



STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number: 151975

TO: Amy H Bowman
Location: REM-2C31&2C18
Art Unit: 1635
Wednesday, May 04, 2005

Case Serial Number: 10/605498

From: Paul Schulwitz
Location: Biotech-Chem Library
REM-1A65
Phone: 571-272-2527

paul.schulwitz@uspto.gov

Search Notes

This Page Blank (uspto)

151975

From: Bowman, Amy
Sent: Thursday, April 28, 2005 12:54 PM
To: STIC-Biotech/ChemLib
Cc: Bowman, Amy
Subject: 10/605,498

Hello,
I need a search of SEQ ID NO: 82 in application 10/605,498, length limited to 20 nucleobases. (I do not need any complements, just matches).
Thanks,
Amy Bowman
AU 1635
REM 2C18

ORFE

STAFF USE ONLY

Searcher: _____
Searcher Phone: 2-_____
Date Searcher Picked up: _____
Date Completed: _____
Searcher Prep/Rev. Time: _____
Online Time: _____

Type of Search

NA#: _____ AA#: _____
Interference: _____ SPDI: _____
S/L: _____ Oligomer: _____
Encode/Transl: _____
Structure#: _____ Text: _____
Inventor: _____ Litigation: _____

Vendors and cost where applicable

STN: _____
DIALOG: _____
QUESTEL/ORBIT: _____
LEXIS/NEXIS: _____
SEQUENCE SYSTEM: _____
WWW/Internet: _____
Other(Specify): _____

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 3, 2005, 17:57:30 ; Search time 124 Seconds
(without alignments)
263.916 Million cell updates/sec

Title: US-10-605-498-82

Perfect score: 20

Sequence: 1 gggacgcgcgtcggtcat 20

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 401682

Minimum DB seq length: 0
Maximum DB seq length: 20

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents_NA:
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	13	65.0	20	3	US-09-046-858A-19
2	13	65.0	20	4	US-09-450-515-19
3	12.4	62.0	20	1	US-08-182-175A-9
4	12.4	62.0	20	1	US-08-474-633A-16
5	12.4	62.0	20	3	US-08-823-771-16
6	12.4	62.0	20	5	PCT-US92-06412-9
7	11.8	59.0	17	1	US-08-171-299B-6
8	11.8	59.0	20	3	US-09-467-642-13
9	11.6	58.0	18	4	US-09-402-618B-75
10	11.4	57.0	17	1	US-08-656-716-18
11	11.4	57.0	17	1	US-08-656-716-42
12	11.4	57.0	17	3	US-08-835-728D-18
13	11.4	57.0	17	3	US-08-835-728D-122
14	11.4	57.0	17	3	US-09-123-951-18
15	11.4	57.0	17	3	US-09-123-951-42
16	11.4	57.0	17	3	US-09-490-558-18
17	11.4	57.0	17	3	US-09-490-558-122
18	11	55.0	15	3	US-09-081-646-391
19	11	55.0	20	3	US-09-702-246-28
20	10.8	54.0	16	4	US-09-479-005A-5
21	10.8	54.0	17	1	US-08-171-299B-5
22	10.8	54.0	17	4	US-09-897-412-7
23	10.8	54.0	18	3	US-09-143-212-58
24	10.8	54.0	20	1	US-08-249-386A-15
25	10.8	54.0	20	2	US-08-478-178A-66
26	10.8	54.0	20	2	US-08-488-177-66
27	10.8	54.0	20	2	US-08-481-072A-66

28	10.8	54.0	20	2	US-08-664-336-66	Sequence 66, Appl
29	10.8	54.0	20	3	US-08-481-066A-66	Sequence 66, Appl
30	10.8	54.0	20	3	US-09-313-930-6	Sequence 6, Appl
31	10.8	54.0	20	3	US-08-829-637A-66	Sequence 66, Appl
32	10.8	54.0	20	4	US-10-025-139-66	Sequence 66, Appl
33	10.8	54.0	20	5	PCT-US95-06160-15	Sequence 15, Appl
34	10.6	53.0	18	3	US-09-218-979-31	Sequence 31, Appl
35	10.6	53.0	18	4	US-09-679-427-31	Sequence 31, Appl
36	10.6	53.0	18	4	US-09-402-618B-77	Sequence 77, Appl
37	10.6	53.0	18	4	US-09-402-618B-79	Sequence 79, Appl
38	10.6	53.0	18	4	US-09-220-536A-31	Sequence 31, Appl
39	10.6	53.0	20	1	US-08-050-073-173	Sequence 173, App
40	10.6	53.0	20	3	US-08-974-549A-477	Sequence 477, App
41	10.6	53.0	20	3	US-09-428-583-13	Sequence 13, Appl
42	10.6	53.0	20	4	US-08-912-951-244	Sequence 244, App
43	10.6	53.0	20	4	US-09-402-181B-477	Sequence 477, App
44	10.6	53.0	20	4	US-09-721-456-477	Sequence 477, App
45	10.4	52.0	14	1	US-08-373-124A-2	Sequence 2, Appl

ALIGNMENTS

RESULT 1
US-09-046-858A-19
; Sequence 19, Application US/09046858A
; Patent No. 6048973
; GENERAL INFORMATION:
; APPLICANT: Rodriguez, Raymond L.
; TITLE OF INVENTION: SUGAR-REGULATORY SEQUENCES
; TITLE OF INVENTION: IN ALPHA-AMYLASE GENES
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: PO Box 60850
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/046,858A
; FILING DATE: 24-MAR-1998
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/042,376
; FILING DATE: 24-MAR-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Dehlinger, Peter J.
; REGISTRATION NUMBER: 28,006
; REFERENCE/DOCKET NUMBER: 2000-0456.30
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-324-0880
; TELEFAX: 650-324-0960
; TELEX:
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-09-046-858A-19

Query Match 65.0%; Score 13; DB 3; Length 20;
Best Local Similarity 100.0%; Pred. No. 7.1e+03;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3 GACGCGGCGTCG 15
|||||||

Db 2 GACGCGGCGTCG 14

RESULT 2

US-09-450-515-19

; Sequence 19, Application US/09450515

; Patent No. 6680425

; GENERAL INFORMATION:

APPLICANT: Rodriguez, Raymond L.

TITLE OF INVENTION: SUGAR-REGULATORY SEQUENCES

IN ALPHA-AMYLASE GENES

NUMBER OF SEQUENCES: 21

CORRESPONDENCE ADDRESS:

ADDRESSEE: Dehlinger & Associates

STREET: PO Box 60850

CITY: Palo Alto

STATE: CA

COUNTRY: USA

ZIP: 94306

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FastSeq for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/450,515

FILING DATE: 29-NO. 6680425-1999

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 09/046,858

FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Dehlinger, Peter J.

REGISTRATION NUMBER: 28,006

REFERENCE/DOCKET NUMBER: 2000-0456.30

TELECOMMUNICATION INFORMATION:

TELEPHONE: 650-324-0880

TELEFAX: 650-324-0960

TELEX: <Unknown>

INFORMATION FOR SEQ ID NO: 19:

SEQUENCE CHARACTERISTICS:

LENGTH: 20 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

SEQUENCE DESCRIPTION: SEQ ID NO: 19:

US-09-450-515-19

Query Match 65.0%; Score 13; DB 4; Length 20;

Best Local Similarity 100.0%; Pred. No. 7.1e+03;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GACGCGGCGTCG 15

|||||

Db 2 GACGCGGCGTCG 14

RESULT 3

US-08-182-175A-9

; Sequence 9, Application US/08182175A

; Patent No. 5559223

; GENERAL INFORMATION:

APPLICANT: Saverio Carl Falco

APPLICANT: Sharon J. Keeler

APPLICANT: Janet A. Rice

TITLE OF INVENTION: Synthetic Storage Proteins with Defined Structure Containing F

NUMBER OF SEQUENCES: 113

CORRESPONDENCE ADDRESS:

ADDRESSEE: E.I. du Pont de Nemours and Company

STREET: 1007 Market Street

CITY: Wilmington

STATE: Delaware

COUNTRY: USA

ZIP: 19898

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy Disk

COMPUTER: Macintosh

OPERATING SYSTEM: Macintosh System, 6.0

SOFTWARE: Microsoft Word, 4.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/182,175A

FILING DATE:

CLASSIFICATION: 800

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 07/743,006

FILING DATE: 9 August 1991

ATTORNEY/AGENT INFORMATION:

NAME: Linda Axamethy Floyd

REGISTRATION NUMBER: 33,692

REFERENCE/DOCKET NUMBER: BB-1031

TELECOMMUNICATION INFORMATION:

TELEPHONE: (302) 992-4929

TELEFAX: (302) 892-7949

TELEX: 835420

INFORMATION FOR SEQ ID NO: 9:

SEQUENCE CHARACTERISTICS:

LENGTH: 20 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

FEATURE:

NAME/KEY: misc feature

LOCATION: 1..20

OTHER INFORMATION: /product= "synthetic oligonucleotide"

OTHER INFORMATION: /standard_name= "SM 70"

US-08-182-175A-9

Query Match 62.0%; Score 12.4; DB 1; Length 20;

Best Local Similarity 92.9%; Pred. NO. 1.3e+04;

Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 CGCGCGCTCGTC 18

|||||

Db 7 CGCTGCGCTCGTC 20

RESULT 4

US-08-474-633A-16

; Sequence 16, Application US/08474633A

; Patent No. 5773691

; GENERAL INFORMATION:

APPLICANT: E. I. DU PONT DE NEMOURS AND

APPLICANT: COMPANY

TITLE OF INVENTION: CHIMERIC GENES AND

TITLE OF INVENTION: METHODS FOR INCREASING

TITLE OF INVENTION: INCREASING THE LYSINE

TITLE OF INVENTION: AND THREONINE CONTENT

TITLE OF INVENTION: OF THE SEEDS OF PLANTS

NUMBER OF SEQUENCES: 107

CORRESPONDENCE ADDRESS:

ADDRESSEE: E. I. DU PONT DE NEMOURS

ADDRESSEE: AND COMPANY

STREET: 1007 MARKET STREET

CITY: WILMINGTON

STATE: DELAWARE

COUNTRY: U.S.A.

ZIP: 19898

COMPUTER READABLE FORM:

MEDIUM TYPE: FLOPPY DISK

COMPUTER: IBM PC COMPATIBLE

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: MICROSOFT WORD VERSION 2.0C

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/474,633A

FILING DATE:

```
/ CLASSIFICATION: 800
/ ATTORNEY/AGENT INFORMATION:
/ NAME: BARBARA C. SIEGELL
/ REGISTRATION NUMBER: 30,684
/ REFERENCE/DOCKET NUMBER: BB-1037-C
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 302-992-4931
/ TELEFAX: 302-773-0164
/ TELEX: 835420
/ INFORMATION FOR SEQ ID NO: 16:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 20 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: DNA (genomic)
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 1..20
/ OTHER INFORMATION: /product= "synthetic
/ oligonucleotide"
/ OTHER INFORMATION: /standard_name= "SM
/ 70"
/
/ US-08-474-633A-16
/
/ Query Match 62.0%; Score 12.4; DB 1; Length 20;
/ Best Local Similarity 92.9%; Pred. No. 1.3e+04;
/ Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
/
/ QY 5 CGCGCGCTCGGTC 18
/ DB 7 CGCTCGCTCGGTC 20
/
/ RESULT 5
/ US-08-823-771-16
/ Sequence 16, Application US/08823771
/ Patent No. 6459019
/ GENERAL INFORMATION:
/ APPLICANT: E. I. DU PONT DE NEMOURS AND
/ COMPANY
/ TITLE OF INVENTION: CHIMERIC GENES AND
/ METHODS FOR INCREASING
/ INCREASING THE LYSINE
/ AND THREONINE CONTENT
/
/ NUMBER OF SEQUENCES: 107
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: E. I. DU PONT DE NEMOURS
/ AND COMPANY
/ STREET: 1007 MARKET STREET
/ CITY: WILMINGTON
/ STATE: DELAWARE
/ COUNTRY: U.S.A.
/ ZIP: 19898
/
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: FLOPPY DISK
/ COMPUTER: IBM PC COMPATIBLE
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: MICROSOFT WORD VERSION 2.0C
/
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/823,771
/ FILING DATE: 24-Mar-1997
/ CLASSIFICATION: <Unknown>
/
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 08/474,633
/ FILING DATE: <Unknown>
/
/ ATTORNEY/AGENT INFORMATION:
/ NAME: BARBARA C. SIEGELL
/ REGISTRATION NUMBER: 30,684
/ REFERENCE/DOCKET NUMBER: BB-1037-C
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 302-992-4931
/ TELEFAX: 302-773-0164
```

```
/ TELEX: 835420
/ INFORMATION FOR SEQ ID NO: 16:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 20 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: DNA (genomic)
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 1..20
/ OTHER INFORMATION: /product= "synthetic
/ oligonucleotide"
/ OTHER INFORMATION: /standard_name= "SM
/ 70"
/
/ US-08-823-771-16
/
/ Query Match 62.0%; Score 12.4; DB 3; Length 20;
/ Best Local Similarity 92.9%; Pred. No. 1.3e+04;
/ Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
/
/ QY 5 CGCGCGCTCGGTC 18
/ DB 7 CGCTCGCTCGGTC 20
/
/ RESULT 6
/ PCT-US92-06412-9
/ Sequence 9, Application PC/TUS9206412
/ GENERAL INFORMATION:
/ APPLICANT: Saverio Carl Falco
/ APPLICANT: Sharon J. Keeler
/ APPLICANT: Janet A. Rice
/ TITLE OF INVENTION: Synthetic Storage Proteins with Defined Structure Containing P
/ NUMBER OF SEQUENCES: 113
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: E.I. du Pont de Nemours and Company
/ STREET: 1007 Market Street
/ CITY: Wilmington
/ STATE: Delaware
/ COUNTRY: USA
/ ZIP: 19898
/
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy Disk
/ COMPUTER: Macintosh
/ OPERATING SYSTEM: Macintosh System, 6.0
/ SOFTWARE: Microsoft Word, 4.0
/
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: PCT/US92/06412
/ FILING DATE: 19920807
/ CLASSIFICATION: 530
/
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 07/743,006
/ FILING DATE: 9-August 1991
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Linda Axamethy Floyd
/ REGISTRATION NUMBER: 33,692
/ REFERENCE/DOCKET NUMBER: BB-1031
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (302) 992-4929
/ TELEFAX: (302) 892-7949
/ TELEX: 835420
/
/ INFORMATION FOR SEQ ID NO: 9:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 20 base pairs
/ TYPE: NUCLEIC ACID
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: DNA (genomic)
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 1..20
```

```
; OTHER INFORMATION: /product= "synthetic oligonucleotide"
; OTHER INFORMATION: /standard_name= "SM 70"
PCT-US92-06412-9

Query Match 62.0%; Score 12.4; DB 5; Length 20;
Best Local Similarity 92.9%; Pred. No. 1.3e+04;
Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 CGCGCGGCTCGGTC 18
DB 7 CGCTCGGCTCGGTC 20

RESULT 7
US-08-171-299B-6/c
; Sequence 6, Application US/08171299B
; Patent No. 5599665
; GENERAL INFORMATION:
; APPLICANT: Barbieri, Joseph T.
; APPLICANT: Frank, Dara W.
; TITLE OF INVENTION: EXOZYME S PROTEIN PREPARATION AND CLONED
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESS: QUARLES & BRADY
; STREET: 411 East Wisconsin Avenue
; CITY: Milwaukee
; STATE: Wisconsin
; COUNTRY: U.S.A.
; ZIP: 53202-4497
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/171,299B
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Baker, Jean C.
; REGISTRATION NUMBER: 35,433
; REFERENCE/DOCKET NUMBER: 650053.90871
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (414) 277-5709
; TELEFAX: (414) 271-3552
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 17 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Other Nucleic Acid
US-08-171-299B-6

Query Match 59.0%; Score 11.8; DB 1; Length 17;
Best Local Similarity 66.7%; Pred. No. 2.5e+04;
Matches 10; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 2 GGACGGCGGCTCGG 16
DB 15 RGAGCGGCGCTSGG 1

RESULT 8
US-09-467-642-13/c
; Sequence 13, Application US/09467642
; Patent No. 6300132
; GENERAL INFORMATION:
; APPLICANT: Brett P. Monia
; APPLICANT: Lex M. Cowsett
; TITLE OF INVENTION: ANTISENSE MODULATION OF TELOMERIC REPEAT BINDING FACTOR 2 EXPRES
; FILE REFERENCE: RTS-0106
; CURRENT APPLICATION NUMBER: US/09/467,642
```

```
; CURRENT FILING DATE: 1999-12-20
; NUMBER OF SEQ ID NOS: 89
; SEQ ID NO 13
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Antisense Oligonucleotide
US-09-467-642-13

Query Match 59.0%; Score 11.8; DB 3; Length 20;
Best Local Similarity 86.7%; Pred. No. 2.5e+04;
Matches 13; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 5 CGCGCGGCTCGGTC 19
DB 18 CGCGCGGATCGGACA 4

RESULT 9
US-09-402-618B-75/c
; Sequence 75, Application US/09402618B
; Patent No. 6709815
; GENERAL INFORMATION:
; APPLICANT: Dong, Fang
; APPLICANT: Lyamichev, Victor
; APPLICANT: Prudent, James
; APPLICANT: Fors, Lance
; APPLICANT: Neri, Bruce
; APPLICANT: Brow, Mary Ann
; APPLICANT: Anderson, Todd
; APPLICANT: Dahlberg, James
; TITLE OF INVENTION: Target-Dependent Reactions Using Structure-Bridging Oligonucleotic
; FILE REFERENCE: FORS-04012
; CURRENT APPLICATION NUMBER: US/09/402,618B
; CURRENT FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: PCT/US98/03194
; PRIOR FILING DATE: 1998-05-05
; NUMBER OF SEQ ID NOS: 128
; SOFTWARE: Patent In version 3.0
; SEQ ID NO 75
; LENGTH: 18
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
US-09-402-618B-75

Query Match 58.0%; Score 11.6; DB 4; Length 18;
Best Local Similarity 77.8%; Pred. No. 3e+04;
Matches 14; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 GGACGGCGGCTCGGTC 19
DB 18 GGGCGCGGCGGCTGTC 1

RESULT 10
US-08-656-716-18
; Sequence 18, Application US/08656716
; Patent No. 5786146
; GENERAL INFORMATION:
; APPLICANT: Herman, James G.
; APPLICANT: Baylin, Stephen B.
; TITLE OF INVENTION: METHYLATION SPECIFIC DETECTION
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
```


; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICANT: US/08/656,716
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Haile, Lisa A.
; REGISTRATION NUMBER: 38,347
; REFERENCE/DOCKET NUMBER: 07265/089001
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 17 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
US-08-656-716-18

Query Match 57.0%; Score 11.4; DB 1; Length 17;
Best Local Similarity 92.3%; Pred. No. 3.7e+04;
Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 5 CGCGGCGCTCGGT 17
Db 3 CGCGGCGTTCGGT 15

RESULT 11
US-08-656-716-42/c
; Sequence 42, Application US/08656716
; Patent No. 5786146
; GENERAL INFORMATION:
; APPLICANT: Herman, James G.
; APPLICANT: Baylin, Stephen B.
; TITLE OF INVENTION: METHYLATION SPECIFIC DETECTION
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/656,716
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Haile, Lisa A.
; REGISTRATION NUMBER: 38,347
; REFERENCE/DOCKET NUMBER: 07265/089001
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 42:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 17 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single

; TOPOLOGY: linear
; MOLECULE TYPE: DNA
US-08-656-716-42

Query Match 57.0%; Score 11.4; DB 1; Length 17;
Best Local Similarity 92.3%; Pred. No. 3.7e+04;
Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 5 CGCGGCGCTCGGT 17
Db 15 CGCGGCGTTCGGT 3

RESULT 12
US-08-835-728D-18
; Sequence 18, Application US/08835728D
; Patent No. 6017704
; GENERAL INFORMATION:
; APPLICANT: Herman, James G.
; APPLICANT: Baylin, Stephen B.
; TITLE OF INVENTION: Methylation Specific Detection
; NUMBER OF SEQUENCES: 216
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/835,728D
; FILING DATE: April 11, 1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/656,716
; FILING DATE: June 03, 1996,
; ATTORNEY/AGENT INFORMATION:
; NAME: Haile, Lisa A.
; REGISTRATION NUMBER: 38,347
; REFERENCE/DOCKET NUMBER: 07265/125001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619/678-5070
; TELEFAX: 619/678-5099
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 17 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
US-08-835-728D-18

Query Match 57.0%; Score 11.4; DB 3; Length 17;
Best Local Similarity 92.3%; Pred. No. 3.7e+04;
Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 5 CGCGGCGCTCGGT 17
Db 3 CGCGGCGTTCGGT 15

RESULT 13
US-08-835-728D-122/c
; Sequence 122, Application US/08835728D
; Patent No. 6017704
; GENERAL INFORMATION:
; APPLICANT: Herman, James G.
; APPLICANT: Baylin, Stephen B.

;; TITLE OF INVENTION: Methylation Specific Detection
;; NUMBER OF SEQUENCES: 216
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Fish & Richardson P.C.
;; STREET: 4225 Executive Square, Suite 1400
;; CITY: La Jolla
;; STATE: CA
;; COUNTRY: USA
;; ZIP: 92037
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/835,728D
;; FILING DATE: April 11, 1997
;; CLASSIFICATION: 435
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 08/656,716
;; FILING DATE: June 03, 1996,
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Haile, Lisa A.
;; REGISTRATION NUMBER: 38,347
;; REFERENCE/DOCKET NUMBER: 07265/125001
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 619/678-5070
;; TELEFAX: 619/678-5099
;; INFORMATION FOR SEQ ID NO: 122:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 17 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: DNA
US-08-835-728D-122

Query Match 57.0%; Score 11.4; DB 3; Length 17;
Best Local Similarity 92.3%; Pred. No. 3.7e+04;
Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 CGCGGCGTTCGGT 17
DB 15 CGCGGCGTTCGGT 3

RESULT 14
US-09-123-951-18
; Sequence 18, Application US/09123951
; Patent No. 6200756
; GENERAL INFORMATION:
; APPLICANT: Herman, James G.
; APPLICANT: Baylin, Stephen B.
; TITLE OF INVENTION: METHYLATION SPECIFIC DETECTION
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/123,951
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/656,716

;; FILING DATE:
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Haile, Lisa A.
;; REGISTRATION NUMBER: 38,347
;; REFERENCE/DOCKET NUMBER: 07265/089001
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 617/542-5070
;; TELEFAX: 617/542-8906
;; TELEX: 200154
;; INFORMATION FOR SEQ ID NO: 18:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 17 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: DNA
US-09-123-951-18

Query Match 57.0%; Score 11.4; DB 3; Length 17;
Best Local Similarity 92.3%; Pred. No. 3.7e+04;
Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 CGCGGCGTTCGGT 17
DB 3 CGCGGCGTTCGGT 15

RESULT 15
US-09-123-951-42/c
; Sequence 42, Application US/09123951
; Patent No. 6200756
; GENERAL INFORMATION:
; APPLICANT: Herman, James G.
; APPLICANT: Baylin, Stephen B.
; TITLE OF INVENTION: METHYLATION SPECIFIC DETECTION
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/123,951
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/656,716
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Haile, Lisa A.
; REGISTRATION NUMBER: 38,347
; REFERENCE/DOCKET NUMBER: 07265/089001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 42:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 17 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
US-09-123-951-42

Query Match 57.0%; Score 11.4; DB 3; Length 17;

Best Local Similarity 92.3%; Pred. No. 3.7e+04;
Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 5 CGCGGGCGCTCGGT 17
Db 15 CGCGGGCGTTCGGT 3

Search completed: May 3, 2005, 21:11:30
Job time : 126 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 3, 2005, 17:44:01 ; Search time 3014 Seconds
(without alignments)
252.583 Million cell updates/sec

Title: US-10-605-498-82

Perfect score: 20

Sequence: 1 gggacggcgctcggtcat 20

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 12452

Minimum DB seq length: 0
Maximum DB seq length: 20

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*

- 1: gb_est1.*
- 2: gb_est2.*
- 3: gb_hic.*
- 4: gb_est3.*
- 5: gb_est4.*
- 6: gb_est5.*
- 7: gb_est6.*
- 8: gb_ges1.*
- 9: gb_ges2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	9.4	47.0	20	8	AZ961140
2	9.2	46.0	19	4	BM395903
3	9.2	46.0	20	1	AJ666384
4	9	45.0	19	8	AZ775541
5	9	45.0	20	7	CF281215
6	8.8	44.0	17	7	CF305567
7	8.8	44.0	18	8	BZ425195
8	8.8	44.0	19	1	AJ649246
9	8.8	44.0	19	8	AZ611716
10	8.8	44.0	20	8	AZ591658
11	8.6	43.0	20	4	BF966452
12	8.6	43.0	20	8	AZ992240
13	8.4	42.0	13	9	AJ594923
14	8.4	42.0	19	4	BM396331
15	8.4	42.0	19	7	CF542982
16	8.4	42.0	19	8	AZ610451
17	8.4	42.0	20	1	AJ801124
18	8.4	42.0	20	4	BM399952
19	8.4	42.0	20	8	AZ822878
20	8.2	41.0	14	5	BQ595292
21	8.2	41.0	15	4	BM399662
22	8.2	41.0	17	4	BM400706
23	8.2	41.0	18	4	BM400305
24	8.2	41.0	18	4	BM400816

C 25	8.2	41.0	19	8	AZ314110	AZ314110 1M0030816
C 26	8.2	41.0	19	8	AZ864551	AZ864551 2M0174M11
C 27	8.2	41.0	20	4	BM401265	BM401265 5009-0-85
C 28	8.2	41.0	20	8	AZ591658	AZ591658 1M0401P19
C 29	8	40.0	14	2	BE516032	BE516032 WHE0629 B
C 30	8	40.0	19	1	AL042746	AL042746 DXF2p437C
C 31	8	40.0	19	6	CA967213	CA967213 CCL01a03
C 32	8	40.0	19	8	AZ849133	AZ849133 2M0150L21
C 33	8	40.0	20	2	AW250737	AW250737 2822626-5
C 34	8	40.0	20	7	CF302285	CF302285 7LEAF--07
C 35	8	40.0	20	8	AZ308410	AZ308410 1M0011A24
C 36	8	40.0	20	8	AZ775705	AZ775705 2M0008P11
C 37	8	40.0	20	9	AG194315	AG194315 Fan trogl
C 38	8	40.0	20	9	CL668826	CL668826 PRI0158D
C 39	7.8	39.0	14	1	AJ682145	AJ682145 AJ682145
C 40	7.8	39.0	15	1	AJ682954	AJ682954 AJ682954
C 41	7.8	39.0	15	2	AW246093	AW246093 2821250.5
C 42	7.8	39.0	16	4	BM395715	BM395715 5009-0-10
C 43	7.8	39.0	16	7	CF303743	CF303743 ABF1--03-
C 44	7.8	39.0	18	4	BM394214	BM394214 50072-2-2
C 45	7.8	39.0	18	4	BM394638	BM394638 50072-2-5

ALIGNMENTS

RESULT 1
AZ961140 20 bp DNA linear GSS 27-APR-2001
LOCUS 2M0229P20F Mouse 10kb plasmid UUC2M library Mus musculus genomic
DEFINITION clone UUC2M0229P20 F, genomic survey sequence.
ACCESSION AZ961140
VERSION AZ961140.1 GI:13832367
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 20)
AUTHORS Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamill, C.,
Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T.,
Reilly, M., Rose, R., Rose, R., Stokes, R., Tingey, A., von
Niederhausern, A. and Wright, D., Weiss, R.
TITLE Mouse whole genome scaffolding with paired end reads from 10kb
JOURNAL plasmid inserts
COMMENT Unpublished (2000)
Contact: Robert B. Weiss
University of Utah Genome Center
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: dunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0229 row: P column: 20
Seq primer: CGTTGTAAACGACGCCAGT
Class: plasmid ends
High quality sequence stop: 20.
Location/Qualifiers
1..20
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clones="UUC2M0229P20"
/sex="Female"
/lab_host="E. coli strain XL10-Gold, Tl-resistant, P-"
/clone_lib="Mouse 10kb plasmid UUC2M library"
/note="Vector: PWD42nv; Purified genomic DNA from M.
musculus C57BL/6J (female) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA

Wed May 4 10:05:34 2005

was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent *E. coli* XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

ORIGIN

Query Match 47.0%; Score 9.4; DB 8; Length 20;
Best Local Similarity 90.9%; Pred. No. 7.4e+06;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 6 GCGGCGCTCGG 16

DB 8 GCTGGCGCTCGG 18

RESULT 2

BM395903 19 bp mRNA linear EST 17-JAN-2002
LOCUS 5009-0-13-G11.t.1 Chilcoat/Turkewitz cDNA (large fraction)
DEFINITION Tetrahymena thermophila cDNA, mRNA sequence.

ACCESSION BM395903

VERSION 1 GI:18195956

KEYWORDS EST

SOURCE Tetrahymena thermophila

ORGANISM Tetrahymena thermophila

REFERENCE 1 (bases 1 to 19)

AUTHORS Turkewitz, A.P., Karer, K.M., Jahn, C., Orlas, E., Kirk, K.E., Frankel, J. and Klobutcher, L.

TITLE EST from Tetrahymena thermophila, strain CU428.1, growing cells

JOURNAL Unpublished (2002)

COMMENT Contact: Turkewitz AP

Molecular Genetics and Cell Biology

University of Chicago

920 E. 58th Street, Chicago, IL 60637, USA

Tel: 773 702 4374

Fax: 773 702 3172

Email: apturkew@midway.uchicago.edu

Seq primer: T3.

Location/Qualifiers

1. 19

/organism="Tetrahymena thermophila"

/mol_type="mRNA"

/strain="CU428.1"

/db_xref="taxon:5911"

/clone_lib="Chilcoat/Turkewitz cDNA (large fraction)"

/note="Vector: Bluescript2 SK+, details on library

preparation can be found in Chilcoat and Turkewitz (2001)

Proc. Natl. Acad. Sci USA, 98: 8709-8713."

ORIGIN

Query Match 46.0%; Score 9.2; DB 4; Length 19;
Best Local Similarity 78.6%; Pred. No. 9.1e+06;
Matches 11; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 ACGCGCGCTCGGT 17

DB 1 ACGCGCGCGCGGT 14

RESULT 3

AJ666384
LOCUS
DEFINITION

20 bp mRNA linear EST 28-JUN-2004
AJ666384 CSEQRAN09 Sus scrofa cDNA clone C0000033_E03, mRNA

ACCESSION AJ666384

VERSION 1 GI:49350835

KEYWORDS EST

SOURCE Sus scrofa (pig)

ORGANISM Sus scrofa

REFERENCE 1 (bases 1 to 20)

AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

TITLE Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.

JOURNAL Development of cDNA and EST resources for studying reproduction and

COMMENT embryo development in pigs and cattle

Unpublished (2004)

Contact: Anderson SI

Genomics and Bioinformatics

Roslin Institute

Roslin, Midlothian, EH25 9PS, UNITED KINGDOM

Single pass sequencing. Bases called and trimmed with phred

v0.020425.c. Vector identified by cross match with the -minscore 20

and -minmatch 12 options. Vector: pBluescriptII (KS+) R. Site 1.

ECORI R. Site 2: NotI Description: Normalised library constructed

from pooled tissue from day 30 placentas. Clones available from UK

Centre for Functional Genomics in Farm Animals, Roslin Institute,

Roslin, Midlothian, UK, EH25 9PS, www.arxgenomics.org.

Location/Qualifiers

1. 20

/organism="Sus scrofa"

/mol_type="mRNA"

/db_xref="taxon:9823"

/clone="C0000033_E03"

/tissue_type="placenta"

/clone_lib="CSEQRAN09"

/note="Vector: pBluescriptII (KS+); Site 1: EORI; Site 2:

NotI; Single pass sequencing. Normalised library

constructed from pooled tissue from day 30 placentas."

ORIGIN

Query Match 46.0%; Score 9.2; DB 1; Length 20;

Best Local Similarity 78.6%; Pred. No. 9.1e+06;

Matches 11; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 GGACGCGCGCTCG 15

DB 2 GGACCGCGTGCA 15

RESULT 4

AZ775541
LOCUS

19 bp DNA linear GSS 16-FEB-2001
DEFINITION 2M0008H17F Mouse 10kb plasmid UUC1M library Mus musculus genomic

ACCESSION AZ775541

VERSION 1 GI:12902185

KEYWORDS GSS

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 19)

AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamil, C.,

Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T.,

Reilly, M., Rose, M., Rose, R., Stokes, R., Tingey, A., von

Niederhausern, A. and Wright, D., Weiss, R.

Mouse whole genome scaffolding with paired end reads from 10kb

plasmid inserts

Unpublished (2000)

Contact: Robert B. Weiss

University of Utah

Genome Center

Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT

84112, USA

Tel: 801 585 5606

Fax: 801 585 7177

Email: ddunn@genetics.utah.edu

Insert Length: 10000 Std Error: 0.00

Plate: 0008 Row: H Column: 17

Seq primer: CGTTGTAAACGACGGCCAGT

Class: plasmid ends

High quality sequence stop: 19.

FEATURES

source

1. 19

Location/Qualifiers

/organism="Mus musculus"

/mol_type="genomic DNA"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/clone="UUGC2M0008H17"

/sex="Male"

/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"

/clone_lib="Mouse 10kb plasmid UUGC1M library"

/note="Vector: PWD42nv; Purified genomic DNA from M.

musculus C57BL/6J (male) was obtained from the Jackson

Laboratory Mouse DNA Resource

(http://www.jax.org/resources/documents/dnares/). The DNA

was hydrodynamically sheared by repeated passage through a

0.005 inch orifice at constant velocity. The sheared DNA

was blunt end-repaired with T4 DNA polymerase and T4

polynucleotide kinase. Adaptor oligonucleotides were

ligated to the blunt ends in high molar excess. The

adapted DNA was purified and size-selected for a 9.5 to

10.5 kb range using preparative agarose gel

electrophoresis. Vector DNA was prepared from a derivative

of pWD42 (gi|4732114|gb|AF129072.1), a copy-number

inducible derivative of plasmid R1. The vector was ligated

with adaptors complementary to the insert adaptors and

purified. The sheared, adapted mouse DNA was annealed to

adapted vector DNA, and transformed into

chemically-competent E. coli XL10-Gold (Stratagene) cells

and selected for ampicillin resistance."

ORIGIN

Query Match 45.0%; Score 9; DB 8; Length 19;

Best Local Similarity 70.6%; Pred. No. 1.1e+07;

Matches 12; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Oy 1 GGGACGGCGGCTCGGT 17

|||||

Db 1 GGGGGGGGGGCTCTGT 17

RESULT 5

CF281215/c

LOCUS

DEFINITION CF281215.1 20 bp mRNA linear EST 14-AUG-2003

14ETL--08-C23.g1 Rice etiolated leaf plasmid cDNA library (14ETL)

Oryza sativa (japonica cultivar-group) cDNA clone 14ETL--08-C23,

mRNA sequence.

ACCESSION CF281215

VERSION CF281215.1 GI:33658602

KEYWORDS EST.

SOURCE Oryza sativa (japonica cultivar-group)

Oryza sativa (japonica cultivar-group)

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;

Ehrhartoideae; Oryzaceae; Oryza.

1 (bases 1 to 20)

Kim,J.S., Jun,K.M., Cheong,P.J., Kim,M.J., Lee,T.H., Shin,Y.C.,

Song,S.I., Kim,J.K., Kim,Y.-K. and Nahm,B.H.

Large-scale Sequencing Analysis of Rice ESTs

Unpublished (2003)

Contact: Nahm B.H.

Genomics and Genetics Institute, GreenGene Biotech Inc.; Division

of Bioscience and Bioinformatics, Myongji University

Yongin, Kyeonggi, Korea

Tel: 82 31 330 6193

Fax: 82 31 321 6355

Email: bnhahm@gbio.com, bnhahm@bio.myongji.ac.kr.

Location/Qualifiers

FEATURES

source

1. 20

/organism="Oryza sativa (japonica cultivar-group)"

/mol_type="mRNA"

/cultivar="Nackdong"

/db_xref="taxon:39947"

/clone="14ETL--08-C23"

/tissue_type="leaf"

/dev_stage="14 days after germination"

/lab_host="E.coli DH10B"

/clone_lib="Rice etiolated leaf plasmid cDNA library

(14ETL)"

/note="Vector: pCR4-TOPO; Site 1: EcoRI; mRNA was capped

with oligoribonucleotides and then used as templates for

RT-PCR."

ORIGIN

Query Match 45.0%; Score 9; DB 7; Length 20;

Best Local Similarity 100.0%; Pred. No. 1.1e+07;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 7 CGCGCGCTCG 15

|||||

Db 16 CGCGCGCTCG 8

RESULT 6

CF305567

LOCUS

DEFINITION

CF305567

HDAL--01-B07.g1 OshDACL-overexpressing transgenic rice lambda phage

cDNA library I (HDAL) Oryza sativa (japonica cultivar-group) cDNA

clone HDAL--01-B07, mRNA sequence.

ACCESSION

CF305567

VERSION

CF305567.1 GI:33677328

KEYWORDS EST.

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

17 bp mRNA linear EST 15-AUG-2003

HDAL--01-B07.g1 OshDACL-overexpressing transgenic rice lambda phage

cDNA library I (HDAL) Oryza sativa (japonica cultivar-group) cDNA

clone HDAL--01-B07, mRNA sequence.

CF305567

HDAL--01-B07.g1 OshDACL-overexpressing transgenic rice lambda phage

cDNA library I (HDAL) Oryza sativa (japonica cultivar-group) cDNA

clone HDAL--01-B07, mRNA sequence.

CF305567

HDAL--01-B07.g1 OshDACL-overexpressing transgenic rice lambda phage

cDNA library I (HDAL) Oryza sativa (japonica cultivar-group) cDNA

clone HDAL--01-B07, mRNA sequence.

CF305567

HDAL--01-B07.g1 OshDACL-overexpressing transgenic rice lambda phage

cDNA library I (HDAL) Oryza sativa (japonica cultivar-group) cDNA

clone HDAL--01-B07, mRNA sequence.

CF305567

HDAL--01-B07.g1 OshDACL-overexpressing transgenic rice lambda phage

cDNA library I (HDAL) Oryza sativa (japonica cultivar-group) cDNA

clone HDAL--01-B07, mRNA sequence.

CF305567

HDAL--01-B07.g1 OshDACL-overexpressing transgenic rice lambda phage

cDNA library I (HDAL) Oryza sativa (japonica cultivar-group) cDNA

clone HDAL--01-B07, mRNA sequence.

CF305567

HDAL--01-B07.g1 OshDACL-overexpressing transgenic rice lambda phage

cDNA library I (HDAL) Oryza sativa (japonica cultivar-group) cDNA

clone HDAL--01-B07, mRNA sequence.

CF305567

HDAL--01-B07.g1 OshDACL-overexpressing transgenic rice lambda phage

cDNA library I (HDAL) Oryza sativa (japonica cultivar-group) cDNA

clone HDAL--01-B07, mRNA sequence.

CF305567

HDAL--01-B07.g1 OshDACL-overexpressing transgenic rice lambda phage

cDNA library I (HDAL) Oryza sativa (japonica cultivar-group) cDNA

clone HDAL--01-B07, mRNA sequence.

CF305567

HDAL--01-B07.g1 OshDACL-overexpressing transgenic rice lambda phage

cDNA library I (HDAL) Oryza sativa (japonica cultivar-group) cDNA

clone HDAL--01-B07, mRNA sequence.

CF305567

HDAL--01-B07.g1 OshDACL-overexpressing transgenic rice lambda phage

cDNA library I (HDAL) Oryza sativa (japonica cultivar-group) cDNA

clone HDAL--01-B07, mRNA sequence.

CF305567

HDAL--01-B07.g1 OshDACL-overexpressing transgenic rice lambda phage

cDNA library I (HDAL) Oryza sativa (japonica cultivar-group) cDNA

clone HDAL--01-B07, mRNA sequence.

CF305567

HDAL--01-B07.g1 OshDACL-overexpressing transgenic rice lambda phage

cDNA library I (HDAL) Oryza sativa (japonica cultivar-group) cDNA

```

Query Match      44.0%; Score 8.8; DB 7; Length 17;
Best Local Similarity 83.3%; Pred. No. 1.4e+07;
Matches 10; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 GCACGCGCGCT 13
Db 2 GCACGAGCGCT 13

RESULT 7
BZ425195/c
LOCUS      18 bp      DNA      linear      GSS 13-DEC-2002
DEFINITION 100026548-5020 Aspergillus terreus random genomic DNA clone library
ACCESSION  BZ425195
VERSION     BZ425195.1 GI:26666650
KEYWORDS
SOURCE      GSS.
ORGANISM    Aspergillus terreus
            Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
            Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
REFERENCE   1 (bases 1 to 18)
AUTHORS     Askenazi, M., Driggers, E.M., Holtzman, D.A., Norman, T.C., Iverson, S.,
            Zimmer, D.P., Boers, M.-E., Blomquist, P.R., Martinez, E.J.,
            Monreal, A.W., Feibelman, T.P., Mayorga, M.E., Maxon, M.E., Sykes, K.,
            Tobin, J., Cordero, E., Salama, S.R., Trueheart, J., Royer, J.C. and
            Madden, K.T.
TITLE       Integrating transcriptional and metabolite profiles to direct the
            engineering of lovastatin-producing strains
JOURNAL     Unpublished (2002)
COMMENT     Contact: Zimmer DP
            Microbia, Inc.
            One Kendall Square Building 1400 W, Cambridge, MA 02139, USA
            Tel: 617-621-8322
            Fax: 617-
            Email: dzimmer@microbia.com
            Class: plasmid ends.

FEATURES             source
            Location/Qualifiers
            1..18
            /organism="Aspergillus terreus"
            /mol_type="genomic DNA"
            /strain="ATCC 20542 (A. terreus Thom, anamorph)"
            /db_xref="taxon:33178"
            /lab_host="Escherichia coli"
            /clone_lib="Aspergillus terreus random genomic DNA clone
            library"
            /note="Vector: pZerOTM-2; Site_1: Sau3A; Site_2: BamHI;
            Sau3A genomic fragments ligated into BamHI digested
            pZerOTM-2"

ORIGIN
Query Match      44.0%; Score 8.8; DB 8; Length 18;
Best Local Similarity 83.3%; Pred. No. 1.4e+07;
Matches 10; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 7 CGCGCTCGGTC 18
Db 15 CCGAGCTCGGTC 4

RESULT 8
AJ649246/c
LOCUS      19 bp      mRNA      linear      EST 07-JUL-2004
DEFINITION AJ649246 CSEQRAN19 Sus scrofa cDNA clone C0003271_P09, mRNA
            sequence.
ACCESSION  AJ649246
VERSION     AJ649246.1 GI:49326091
KEYWORDS   EST.
SOURCE     Sus scrofa (pig)
ORGANISM   Sus scrofa
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Euthera; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE   1 (bases 1 to 19)
            /organism="Sus scrofa"
            /mol_type="mRNA"
            /db_xref="taxon:9823"
            /clone="C0003271_P09"
            /tissue_type="ovary"
            /clone_lib="CSEQRAN19"
            /note="Vector: pBluescriptII(KS+); Site_1: EcoRI; Site_2:
            NotI; Single pass sequencing; Normalised library
            constructed from pooled ovaries"

ORIGIN
Query Match      44.0%; Score 8.8; DB 1; Length 19;
Best Local Similarity 83.3%; Pred. No. 1.4e+07;
Matches 10; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 8 GCGCTCGGTCA 19
Db 19 GCGCTCGGTCA 8

RESULT 9
AZ611716
LOCUS      19 bp      DNA      linear      GSS 13-DEC-2000
DEFINITION 1M0438B15F Mouse 10kb plasmid UUC1M library mus musculus genomic
            clone UUGC1M0438B15 F, genomic survey sequence.
ACCESSION  AZ611716
VERSION     AZ611716.1 GI:11733906
KEYWORDS   GSS.
SOURCE     Mus musculus (house mouse)
ORGANISM   Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 19)
AUTHORS     Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamil, C.,
            Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T.,
            Reilly, M., Rose, M., Rose, R., Stokes, R., Tingey, A., von
            Niedernausern, A. and Wright, D., Weiss, R.,
            Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
            Unpublished (2000)
            Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 309, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000 Std Error: 0.00
            Plate: 0438 row: B column: 15
            Seq primer: CCGTGAACGACGCGCAGT
            Class: plasmid ends
            High quality sequence stop: 19.
            Location/Qualifiers
            1..19
            /organism="Mus musculus"

FEATURES             source
            Location/Qualifiers

```

```

ANDERSON, S.I., Finlayson, H.A. and Archibald, A.L.
Development of cDNA and EST resources for studying reproduction and
embryo development in pigs and cattle
Unpublished (2004)
Contact: Anderson SI
Genomics and Bioinformatics
Roslin Institute
Roslin, Midlothian, EH25 9PS, UNITED KINGDOM
Single pass sequencing. Bases called and trimmed with phred
v0.020425.c. Vector identified by cross match with the -minscore 20
and -mismatch 12 options. Vector: pBluescriptII(KS) R. Site1: EcoRI
R. Site2: NotI 5' Seq Primer M13F Normalised library constructed
from pooled ovaries. Clones available from UK Centre for Functional
Genomics in Farm Animals, Roslin Institute, Roslin, Midlothian, UK,
EH25 9PS, www.ark-genomics.org.
Location/Qualifiers
1..19
/organism="Sus scrofa"
/mol_type="mRNA"
/db_xref="taxon:9823"
/clone="C0003271_P09"
/tissue_type="ovary"
/clone_lib="CSEQRAN19"
/note="Vector: pBluescriptII(KS+); Site_1: EcoRI; Site_2:
NotI; Single pass sequencing; Normalised library
constructed from pooled ovaries"

ORIGIN
Query Match      44.0%; Score 8.8; DB 1; Length 19;
Best Local Similarity 83.3%; Pred. No. 1.4e+07;
Matches 10; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 8 GCGCTCGGTCA 19
Db 19 GCGCTCGGTCA 8

RESULT 9
AZ611716
LOCUS      19 bp      DNA      linear      GSS 13-DEC-2000
DEFINITION 1M0438B15F Mouse 10kb plasmid UUC1M library mus musculus genomic
            clone UUGC1M0438B15 F, genomic survey sequence.
ACCESSION  AZ611716
VERSION     AZ611716.1 GI:11733906
KEYWORDS   GSS.
SOURCE     Mus musculus (house mouse)
ORGANISM   Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 19)
AUTHORS     Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamil, C.,
            Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T.,
            Reilly, M., Rose, M., Rose, R., Stokes, R., Tingey, A., von
            Niedernausern, A. and Wright, D., Weiss, R.,
            Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
            Unpublished (2000)
            Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 309, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000 Std Error: 0.00
            Plate: 0438 row: B column: 15
            Seq primer: CCGTGAACGACGCGCAGT
            Class: plasmid ends
            High quality sequence stop: 19.
            Location/Qualifiers
            1..19
            /organism="Mus musculus"

FEATURES             source
            Location/Qualifiers

```



```

/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC1M0438B15"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/clone_lib="Mouse 10kb plasmid UUGC1M library"
/note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

ORIGIN
Query Match 44.0%; Score 8.8; DB 8; Length 19;
Best Local Similarity 83.3%; Pred. No. 1.4e+07;
Matches 10; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 6 GCGGCGCTCGGT 17
    |||||
Db 2 GCGGCGCGGCT 13

RESULT 10
AZ591658/c
LOCUS
DEFINITION
clone UUGC1M0401F19 R, genomic survey sequence.
ACCESSION
AZ591658
VERSION
AZ591658.1 GI:11713848
KEYWORDS
GSS.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 20)
Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamill, C.,
Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T.,
Reilly, M., Rose, M., Rose, R., Stokes, R., Tingey, A., von
Niederhausern, A. and Wright, D., Weiss, R.
Niederhausern, A. and Wright, D., Weiss, R.
Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts
Unpublished (2000)
Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: dduan@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0401 row: F column: 19
Seq primer: CACACAGGAACAGCTATGACC
Class: plasmid ends
High quality sequence stop: 20.
Location/Qualifiers
1..20
/organism="Mus musculus"

FEATURES
source

```

```

/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC1M0401F19"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/clone_lib="Mouse 10kb plasmid UUGC1M library"
/note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

ORIGIN
Query Match 44.0%; Score 8.8; DB 8; Length 20;
Best Local Similarity 83.3%; Pred. No. 1.4e+07;
Matches 10; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 6 GCGGCGCTCGGT 17
    |||||
Db 20 GCAGCGCTGGGT 9

RESULT 11
BF966452
LOCUS
DEFINITION
602287066F1 NIH_MGC_95 Homo sapiens cDNA clone IMAGE:4375648 5',
mRNA sequence.
ACCESSION
BF966452
VERSION
BF966452.2 GI:12388052
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 20)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
On Jan 16, 2001 this sequence version replaced gi:12333667.
Contact: Robert Strausberg, Ph.D.
Email: cgsbbs-remail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
Toshiyuki and Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10041 row: d column: 17
High quality sequence stop: 20.
Location/Qualifiers
1..20
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4375648"
/tissue_type="hippocampus"

FEATURES
source

```

```

/lab_host="DH10B"
/clone_lib="NIH_MGC_95"
/notes="Organ: brain; Vector: pBluescriptR (modified
pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI
(gtcgag); Oligo-dr primed using primer
5'-TTTTTTTTTTTTTNN-3', size-selected for average
insert size 2.5 kb and normalized to ROT 5. This is a
primary library enriched for full-length clones and
constructed using the Cap-trapper method (Carninci, in
preparation). Library constructed by M. Brownstein
(NIMH/NHGRI, National Institutes of Health). Note: this
is a NIH_MGC Library."

ORIGIN
Query Match 43.0%; Score 8.6; DB 4; Length 20;
Best Local Similarity 73.3%; Pred. No. 1.7e+07;
Matches 11; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 GGACGCGCGCTCGG 16
||| ||| ||| |||
Db 4 GGAGTCGGTCGCGG 18

RESULT 12
AZ992240/c
LOCUS 20 bp DNA linear GSS 27-APR-2001
DEFINITION 2M0276M15R Mouse 10kb plasmid UUGC2M library Mus musculus genomic
clone UUGC2M0276M15 R, genomic survey sequence.
ACCESSION AZ992240.1 GI:13863467
VERSION GSS.
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM Mus musculus

REFERENCE
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenan,E., Pedersen,T.,
Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von
Niederhausern,A. and Wright,D.,Weiss,R.
Niederhausern,A. and Wright,D.,Weiss,R.
Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts
Unpublished (2000)
Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0276 row: M column: 15
Seq primer: CACACAGGAACAGTATGACC
Class: plasmid ends
High quality sequence stop: 20.
Location/Qualifiers
1..20
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="CS7BL/6J"
/db_xref="taxon:10090"
/clone="UUGC2M0276M15"
/sex="Female"
/lab_host="E. coli strain XL10-Gold, Tl-resistant, F-"
/clone_lib="Mouse 10kb plasmid UUGC2M library"
/notes="Vector: PWD42nv; Purified genomic DNA from M.
musculus C57BL/6J (female) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4

/clone_lib="NIH_MGC_95"
/notes="Organ: brain; Vector: pBluescriptR (modified
pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI
(gtcgag); Oligo-dr primed using primer
5'-TTTTTTTTTTTTTNN-3', size-selected for average
insert size 2.5 kb and normalized to ROT 5. This is a
primary library enriched for full-length clones and
constructed using the Cap-trapper method (Carninci, in
preparation). Library constructed by M. Brownstein
(NIMH/NHGRI, National Institutes of Health). Note: this
is a NIH_MGC Library."

ORIGIN
Query Match 43.0%; Score 8.6; DB 8; Length 20;
Best Local Similarity 73.3%; Pred. No. 1.7e+07;
Matches 11; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 GGGACGCGCGCTCG 15
||| ||| ||| |||
Db 20 GGTACGCGGTGCTCG 6

RESULT 13
AJ594923/c
LOCUS 13 bp DNA linear GSS 15-JAN-2004
DEFINITION Arabidopsis thaliana T-DNA flanking sequence, left border, clone
408B05, genomic survey sequence.
ACCESSION AJ594923.1 GI:37944547
VERSION GSS; left border; T-DNA flanking sequence.
KEYWORDS Arabidopsis thaliana (thale cress)
SOURCE Arabidopsis thaliana
ORGANISM Arabidopsis thaliana

REFERENCE
AUTHORS Brunaud,V., Balzerque,S., Dubreucq,B., Aubourg,S., Samson,F.,
Chauvin,S., Bechtold,N., Cruaud,C., DeRose,R., Pelletier,G.,
Lepiniec,B., Caboche,M. and Lecharny,A.
T-DNA integration into the Arabidopsis genome depends on sequences
of pre-insertion sites
EMBO Rep. 3 (12), 1152-1157 (2002)
22363535
12446565
2 (bases 1 to 13)
Balzerque,S.
Direct Submission
Submitted (23-OCT-2003) Balzerque S., UMRGV, INRA/CNRS, 2 rue
Gaston Cremieux, 91057 Evry cedex, FRANCE
PCR was performed on DNA from transformants of Arabidopsis thaliana
plants from INRA (Versailles). The DNA fragment (8) resulting from
the PCR were directly sequenced from the left or the right border
to determine the genomic sequence flanking the insertion. T-DNA
derived sequences were removed. Information to order the
corresponding mutant line and a link to a database providing a
graphical display of the insertion site are available at
http://dbgap.versailles.inra.fr/publiclines/. This sequence has
been generated in the framework of the French plant genomics
program 'Genoplante' (http://www.genoplante.com and
http://genoplante-info.infobiogen.fr).
Location/Qualifiers
1..13
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/cultivar="Wassiliewskija"
/db_xref="taxon:3702"
/clone="408B05"
/misc_feature 1..13
/clone_lib="Arabidopsis thaliana T-DNA insertion lines"
/notes="T-DNA flanking sequence
left border"

ORIGIN

```

```
Query Match      42.0%; Score 8.4; DB 9; Length 13;
Best Local Similarity 90.0%; Pred. No. 2.2e+07;
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 10 GCCTCGGTCA 19
    ||| ||| ||| |||
Db 13 CGCGGGTCA 4

RESULT 14
BM396331/C      19 bp mRNA linear EST 17-JAN-2002
LOCUS           5009-0-2-E02.t.1 Chilcoat/Turkewitz cDNA (large fraction)
DEFINITION      Tetrahymena thermophila cDNA, mRNA sequence.
ACCESSION       BM396331
VERSION         BM396331.1 GI:18196384
KEYWORDS        EST.
SOURCE          Tetrahymena thermophila
ORGANISM        Tetrahymena thermophila
Eukaryota; Alveolata; Ciliophora; Oligohymenophorea;
Hymenostomatida; Tetrahymenina; Tetrahymenidae; Tetrahymena.
REFERENCE       1 (bases 1 to 19)
AUTHORS         Turkewitz,A.P., Karrer,K.M., Jahn,C., Orias,E., Kirk,K.E.,
                Frankel,J. and Klobutcher,L.
TITLE           EST from Tetrahymena thermophila, strain CU428.1, growing cells
JOURNAL         Unpublished (2002)
COMMENT         Contact: Turkewitz AP
                Molecular Genetics and Cell Biology
                University of Chicago
                920 E. 58th Street, Chicago, IL 60637, USA
                Tel: 773 702 4374
                Fax: 773 702 3172
                Email: apturkew@midway.uchicago.edu
                Seq primer: T3.

FEATURES         Location/Qualifiers
                source
                1..19
                /organism="Tetrahymena thermophila"
                /mol_type="mRNA"
                /strain="CU428.1"
                /db_xref="taxon:5911"
                /clone_lib="Chilcoat/Turkewitz cDNA (large fraction)"
                /note="Vector: Bluescript2 SK+; Details on library
                preparation can be found in Chilcoat and Turkewitz (2001)
                Proc. Natl. Acad. Sci USA, 98: 8709-8713."

ORIGIN
Query Match      42.0%; Score 8.4; DB 4; Length 19;
Best Local Similarity 66.7%; Pred. NO. 2.1e+07;
Matches 12; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 3 GACGGCGGCTCGGTGCAT 20
    ||| ||| ||| ||| |||
Db 19 GCCACCGCGGTGATGCAT 2

RESULT 15
CF542982        19 bp mRNA linear EST 22-SEP-2003
LOCUS           S014680w-024-030-P12-SP6 MP1Z-ADIS-024-leaf Beta vulgaris cDNA
DEFINITION      clone 024-030-P12 5-PRIME, mRNA sequence.
ACCESSION       CF542982
VERSION         CF542982.1 GI:34891422
KEYWORDS        EST.
SOURCE          Beta vulgaris
ORGANISM        Beta vulgaris
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Caryophyllales; Amaranthaceae; Beta.
REFERENCE       1 (bases 1 to 19)
AUTHORS         Herwig,R.; Schulz,B., Weishaar,B., Hennig,S., Steinfath,M.,
                Drungowski,M., Stahl,D., Wruck,W., Menze,A., O'Brien,J., Lehrach,H.
                and Radelof,U.

TITLE           Construction of a 'unigene' cDNA clone set by oligonucleotide
                fingerprinting allows access to 25 000 potential sugar beet genes
                Plant J. 32 (5), 845-857 (2002)
JOURNAL         22362189
MEDLINE         12472698
PUBMED          Contact: Weishaar B
                ADIS DNA core facility at MP1Z
                Max-Planck-Institute for Plant Breeding Research
                Carl-von-Linne Weg 10, 50829 Koeln, Germany
                Fax: 00492215062851
                Email: weishaar@mpiz-koeln.mpg.de
                Insert Length: 19 Std Error: 0.00
                Plate: 30 row: P column: 12
                Seq primer: SP6.
                Location/Qualifiers
                source
                1..19
                /organism="Beta vulgaris"
                /mol_type="mRNA"
                /cultivar="KWS2320 (double haploid, monogerm breeding
                line)"
                /db_xref="GABI:936786"
                /db_xref="taxon:161934"
                /clone="024-030-P12"
                /tissue_type="leaf"
                /lab_host="EMDH10B"
                /clone_lib="MP1Z-ADIS-024-leaf"
                /note="Vector: pCMVSPORT6; Site 1: SalI; Site 2: NotI;
                cDNA library from sugar beet, library provided by KWS
                Kleinwanzlebener Saatzzucht AG Einbeck, Germany, contact:
                b.schulz@kws.de; cloning sites SalI-NotI, primer sites and
                orientation:
                SP6-Sali-CCACGCGTCCG-5prime-cDNA-polyA-CC-NotI-T7; Note:
                Sequencing granted in the context of the GABI-Beet
                project, local PI: Dr. Katharina Schneider, coordinator:
                Prof. Christian Jung; Sequence submission managed by
                RZPD/GABI-Primary database:http://gabi.rzpd.de"

ORIGIN
Query Match      42.0%; Score 8.4; DB 7; Length 19;
Best Local Similarity 90.0%; Pred. No. 2.1e+07;
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 GGACGCGGCG 11
    ||| ||| |||
Db 2 GGACGCGGCG 11

Search completed: May 3, 2005, 21:09:19
Job time : 3021 secs
```

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 3, 2005, 19:41:30 ; Search time 489 Seconds
(without alignments)
248.985 Million cell updates/sec

Title: US-10-605-498-82

Perfect score: 20

Sequence: 1 gggacgcggcctcggtcat 20

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 5642217 seqs, 3043843248 residues

Total number of hits satisfying chosen parameters: 1563430

Minimum DB seq length: 0

Maximum DB seq length: 20

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:*

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/2/pubpna/PTCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/2/pubpna/PTCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
- 13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
- 14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
- 16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq:*
- 17: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq:*
- 18: /cgn2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq:*
- 19: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
- 20: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:*
- 21: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
- 22: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	20	100.0	20	18	US-10-605-498-82
2	19	95.0	19	18	US-10-605-498-82
3	13	65.0	20	14	US-10-206-654-19
4	12.6	63.0	20	17	US-10-131-827-8773
5	12.4	62.0	20	14	US-10-023-066A-16
6	12.4	62.0	20	18	US-10-804-678-16
7	11.8	59.0	19	15	US-10-198-677-20
8	11.8	59.0	20	17	US-10-148-355A-13
9	11.8	59.0	20	18	US-10-394-808-28
10	11.8	59.0	20	18	US-10-394-808-103
11	11.8	59.0	20	18	US-10-803-482-28

12	11.8	59.0	20	18	US-10-803-482-103
13	11.6	58.0	18	9	US-09-942-588A-31
14	11.6	58.0	18	9	US-09-764-420A-32
15	11.6	58.0	18	9	US-09-942-596A-31
16	11.6	58.0	18	10	US-09-988-873A-31
17	11.6	58.0	18	10	US-09-882-345A-75
18	11.6	58.0	18	10	US-09-942-662A-31
19	11.6	58.0	18	10	US-09-764-420A-32
20	11.6	58.0	18	14	US-10-231-302-31
21	11.6	58.0	18	17	US-10-608-804-31
22	11.6	58.0	18	17	US-10-634-510-31
23	11.6	58.0	18	17	US-10-807-114-75
24	11.6	58.0	18	19	US-10-655-362-75
25	11.6	58.0	20	10	US-09-888-361-147
26	11.4	57.0	13	17	US-10-391-415-20
27	11.2	56.0	16	18	US-10-712-672-1620
28	11.2	56.0	17	16	US-10-230-006-530
29	11.2	56.0	17	16	US-10-230-006-597
30	11.2	56.0	17	16	US-10-230-006-598
31	11.2	56.0	18	11	US-09-825-566-47
32	11.2	56.0	18	18	US-10-240-126-47
33	11.2	56.0	19	18	US-10-670-011-7
34	11.2	56.0	19	18	US-10-670-011-103
35	11.2	56.0	19	19	US-10-764-957-7
36	11.2	56.0	19	19	US-10-764-957-103
37	11.2	56.0	20	10	US-09-972-607-86
38	11.2	56.0	20	17	US-10-628-841-86
39	11.2	56.0	20	18	US-10-619-739-1280
40	11	55.0	19	17	US-10-016-490C-51
41	11	55.0	20	9	US-09-791-243-28
42	11	55.0	20	10	US-09-963-668B-9
43	11	55.0	20	14	US-10-096-595B-4
44	10.8	54.0	15	9	US-09-504-231A-926
45	10.8	54.0	15	9	US-09-274-553D-926

ALIGNMENTS

RESULT 1

US-10-605-498-82
; Sequence 82, Application US/10605498
; Publication No. US20040127441A1
; GENERAL INFORMATION:
; APPLICANT: Rocchi, Palma
; APPLICANT: Signaevsky, Maxim
; TITLE OF INVENTION: Compositions and Methods for Treatment of Prostate and Other
; TITLE OF INVENTION: Cancers
; FILE REFERENCE: UBC.P-031
; CURRENT APPLICATION NUMBER: US/10/605,498
; CURRENT FILING DATE: 2003-10-02
; PRIOR FILING DATE: 2002-10-02
; PRIOR FILING DATE: 2002-10-02
; PRIOR APPLICATION NUMBER: US 60/463,952
; PRIOR FILING DATE: 2003-04-18
; NUMBER OF SEQ ID NOS: 91
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 82
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-605-498-82

Query Match 100.0%; Score 20; DB 18; Length 20;
Best Local Similarity 100.0%; Pred. No. 4.1;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGACGCGCGCTCGGTGTCAT 20

Db 1 GGGACGCGCGCTCGGTGTCAT 20

```
RESULT 2
US-10-605-498-90/c
; Sequence 90, Application US/10605498
; Publication No. US20040127441A1
; GENERAL INFORMATION:
; APPLICANT: Gleave, Martin
; APPLICANT: Rocchi, Palma
; APPLICANT: Signaevsky, Maxim
; TITLE OF INVENTION: Compositions and Methods for Treatment of Prostate and Other
; FILE REFERENCE: UBC-P-031
; CURRENT APPLICATION NUMBER: US/10/605,498
; CURRENT FILING DATE: 2003-10-02
; PRIOR APPLICATION NUMBER: US 60/415,859
; PRIOR FILING DATE: 2002-10-02
; PRIOR APPLICATION NUMBER: US 60/463,952
; PRIOR FILING DATE: 2003-04-18
; NUMBER OF SEQ ID NOS: 91
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 90
; LENGTH: 19
; TYPE: RNA
; ORGANISM: Homo sapiens
US-10-605-498-90

Query Match.          95.0%; Score 19; DB 18; Length 19;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GGACGGCGGCTCGTGCAT 20
DB 19 GGACGGCGGCTCGTGCAT 1

RESULT 3
US-10-206-654-19
; Sequence 19, Application US/10206654
; Publication No. US20030033634A1
; GENERAL INFORMATION:
; APPLICANT: Rodriguez, Raymond L.
; TITLE OF INVENTION: SUGAR-REGULATORY SEQUENCES
; IN ALPHA-AMYLASE GENES
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: PO Box 60850
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/206,654
; FILING DATE: 25-Jul-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/046,858A
; FILING DATE: 24-MAR-1998
; APPLICATION NUMBER: 60/042,376
; FILING DATE: 24-MAR-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Dehlinger, Peter J.
; REGISTRATION NUMBER: 28,006
; REFERENCE/DOCKET NUMBER: 2000-0456.30
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-324-0880
; TELEFAX: 650-324-0960
; TELEX: <Unknown>
; INFORMATION FOR SEQ ID NO: 19:
```

```
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 19:
US-10-206-654-19

Query Match          65.0%; Score 13; DB 14; Length 20;
Best Local Similarity 100.0%; Pred. No. 1.1e+04;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 GACGGCGGCTCG 15
DB 2 GACGGCGGCTCG 14

RESULT 4
US-10-131-827-8773
; Sequence 8773, Application US/10131827
; Publication No. US20040009479A1
; GENERAL INFORMATION:
; APPLICANT: Wohlgenuth, Jay
; APPLICANT: Fry, Kirk
; APPLICANT: Woodward, Robert
; APPLICANT: Ly, Ngoc
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING AND MONITORING AUTOIMMUNE
; DISEASES
; TITLE OF INVENTION: CHRONIC INFLAMMATORY DISEASES
; FILE REFERENCE: 506612000120
; CURRENT APPLICATION NUMBER: US/10/131,827
; CURRENT FILING DATE: 2002-09-06
; PRIOR APPLICATION NUMBER: US 10/006,290
; PRIOR FILING DATE: 2001-10-22
; PRIOR APPLICATION NUMBER: US 60/296,764
; PRIOR FILING DATE: 2001-06-08
; NUMBER OF SEQ ID NOS: 9090
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 8773
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Human cytomegalovirus
US-10-131-827-8773

Query Match          63.0%; Score 12.6; DB 17; Length 20;
Best Local Similarity 78.9%; Pred. No. 1.6e+04;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 GGACGGCGGCTCGTGCAT 20
DB 1 GTACGGCGTCTGGTGCAT 19

RESULT 5
US-10-023-066A-16
; Sequence 16, Application US/10023066A
; Publication No. US20030056242A1
; GENERAL INFORMATION:
; APPLICANT: E. I. DU PONT DE NEMOURS AND
; COMPANY
; TITLE OF INVENTION: CHIMERIC GENES AND METHODS FOR
; INCREASING THE LYSINE AND
; THREONINE CONTENT OF THE SEEDS OF
; PLANTS
; NUMBER OF SEQUENCES: 107
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: E. I. DU PONT DE NEMOURS
; AND COMPANY
; STREET: 1007 MARKET STREET
; CITY: WILMINGTON
; STATE: DELAWARE
; COUNTRY: U.S.A.
; ZIP: 19898
; COMPUTER READABLE FORM:
```

MEDIUM TYPE: FLOPPY DISK
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: MICROSOFT WORD VERSION 2.0C
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/023,066A
FILING DATE: 29-Apr-2002
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: BARBARA C. SIEGELL
REGISTRATION NUMBER: 30,684
REFERENCE/DOCKET NUMBER: BB-1037-C
TELECOMMUNICATION INFORMATION:
TELEPHONE: 302-992-4931
TELEFAX: 302-773-0164
TELEX: 835420
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 20 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: misc_feature
LOCATION: 1..20
OTHER INFORMATION: /product= "synthetic"
oligonucleotide"
/standard_name= "SM"
70"
SEQUENCE DESCRIPTION: SEQ ID NO: 16:
US-10-023-066A-16
Query Match 62.0%; Score 12.4; DB 14; Length 20;
Best Local Similarity 92.9%; Pred. No. 2.1e+04;
Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 5 CGCGCGCTCGGTC 18
DB 7 CGCTGCGCTCGGTC 20
RESULT 6
US-10-804-678-16
Sequence 16, Application US/10804678
Publication No. US20050005330A1
GENERAL INFORMATION:
APPLICANT: EPELBAUM, SABINE URSULA
FALCO, SAVERIO CARL
MCDEVITT, RAYMOND ERVIN, III
TITLE OF INVENTION: CHIMERIC GENES AND METHODS FOR
INCREASING THE LYSINE CONTENT OF
THE SEEDS OF PLANTS
NUMBER OF SEQUENCES: 132
CORRESPONDENCE ADDRESS:
ADDRESSEE: E. I. DU PONT DE NEMOURS AND COMPANY
STREET: 1007 MARKET STREET
CITY: WILMINGTON
STATE: DELAWARE
COUNTRY: U.S.A.
ZIP: 19898
COMPUTER READABLE FORM:
MEDIUM TYPE: DISKETTE, 3.50 INCH
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: MICROSOFT OFFICE 97
SOFTWARE: MICROSOFT WINDOWS 95
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/804,678
FILING DATE: 19-Mar-2004
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/049,304
FILING DATE: 27-Mar-1998

APPLICATION NUMBER: 08/824,627
FILING DATE: MARCH 27, 1997
ATTORNEY/AGENT INFORMATION:
NAME: CHRISTENBURY, LYNNE M.
REGISTRATION NUMBER: 30,971
REFERENCE/DOCKET NUMBER: BB-1037-P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 302-992-5481
TELEFAX: 302-892-7949
TELEX: 835420
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 20 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: misc_feature
LOCATION: 1..20
OTHER INFORMATION: /product= "synthetic"
oligonucleotide"
/standard_name= "SM"
70"
SEQUENCE DESCRIPTION: SEQ ID NO: 16:
US-10-804-678-16
Query Match 62.0%; Score 12.4; DB 18; Length 20;
Best Local Similarity 92.9%; Pred. No. 2.1e+04;
Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 5 CGCGCGCTCGGTC 18
DB 7 CGCTGCGCTCGGTC 20
RESULT 7
US-10-198-677-20
Sequence 20, Application US/10198677
Publication No. US20030119023A1
GENERAL INFORMATION:
APPLICANT: CHOO, Yen
APPLICANT: KLUG, Aaron
APPLICANT: MOORE, Michael
TITLE OF INVENTION: NUCLEIC ACID BINDING POLYPEPTIDES CHARACTERIZED BY
FLEXIBLE LINKERS CONNECTED NUCLEIC ACID BINDING
MOLECULES
TITLE OF INVENTION: MOLECULES
FILE REFERENCE: 8325-2011 / G11-US1
CURRENT APPLICATION NUMBER: US/10/198,677
CURRENT FILING DATE: 2002-07-17
NUMBER OF SEQ ID NOS: 144
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 20
LENGTH: 19
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: bsG
US-10-198-677-20
Query Match 59.0%; Score 11.8; DB 15; Length 19;
Best Local Similarity 86.7%; Pred. No. 4.1e+04;
Matches 13; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 2 GGACGGCGGCTCGG 16
DB 3 GGACGGCGGCTGGG 17
RESULT 8
US-10-148-355A-13/C
Sequence 13, Application US/10148355A
Publication No. US20030207831A1

```
; GENERAL INFORMATION:
; APPLICANT: Brett P. Monia
; APPLICANT: Lex M. Cowert
; APPLICANT: ISIS PHARMACEUTICALS, INC.
; TITLE OF INVENTION: ANTISENSE MODULATION OF TELOMERIC REPEAT BINDING FACTOR 2
; FILE REFERENCE: RTSP-0082
; CURRENT APPLICATION NUMBER: US/10/148,355A
; PRIOR FILING DATE: 2002-09-30
; PRIOR FILING DATE: 1999-12-17
; NUMBER OF SEQ ID NOS: 89
; SEQ ID NO 13
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Antisense Oligonucleotide
US-10-148-355A-13

Query Match          59.0%; Score 11.8; DB 17; Length 20;
Best Local Similarity 86.7%; Pred. No. 4e+04; 2; Indels 0; Gaps 0;
Matches 13; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 5 CGCGGCGCTCGTCA 19
Db 18 CGCGGCGATCGACA 4

RESULT 9
US-10-394-808-28/c
; Sequence 28, Application US/10394808
; Publication No. US2004018559A1
; GENERAL INFORMATION:
; APPLICANT: Brett P. Monia
; APPLICANT: Mark J. Graham
; TITLE OF INVENTION: MODULATION OF DIACYLGLYCEROL ACYLTRANSFERASE 1 EXPRESSION
; FILE REFERENCE: BIOL000303
; CURRENT APPLICATION NUMBER: US/10/394,808
; CURRENT FILING DATE: 2003-03-21
; NUMBER OF SEQ ID NOS: 152
; SEQ ID NO 28
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Antisense Oligonucleotide
US-10-394-808-28

Query Match          59.0%; Score 11.8; DB 18; Length 20;
Best Local Similarity 86.7%; Pred. No. 4e+04; 2; Indels 0; Gaps 0;
Matches 13; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GGGACGCGCGCTCG 15
Db 20 GGGTCGCGGCCCTCG 6

RESULT 10
US-10-394-808-103
; Sequence 103, Application US/10394808
; Publication No. US2004018559A1
; GENERAL INFORMATION:
; APPLICANT: Brett P. Monia
; APPLICANT: Mark J. Graham
; TITLE OF INVENTION: MODULATION OF DIACYLGLYCEROL ACYLTRANSFERASE 1 EXPRESSION
; FILE REFERENCE: BIOL000303
; CURRENT APPLICATION NUMBER: US/10/394,808
; CURRENT FILING DATE: 2003-03-21
; NUMBER OF SEQ ID NOS: 152
; SEQ ID NO 103
; LENGTH: 20
; TYPE: RNA
; ORGANISM: H. sapiens
; FEATURE:
; OTHER INFORMATION: Antisense Oligonucleotide
US-10-394-808-103

Query Match          59.0%; Score 11.8; DB 18; Length 20;
Best Local Similarity 86.7%; Pred. No. 4e+04; 2; Indels 0; Gaps 0;
Matches 13; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GGGACGCGCGCTCG 15
Db 20 GGGTCGCGGCCCTCG 6

RESULT 11
US-10-803-482-28/c
; Sequence 28, Application US/10803482
; Publication No. US20040209838A1
; GENERAL INFORMATION:
; APPLICANT: Brett P. Monia
; APPLICANT: Mark J. Graham
; TITLE OF INVENTION: MODULATION OF DIACYLGLYCEROL ACYLTRANSFERASE 1 EXPRESSION
; FILE REFERENCE: BIOL000303
; CURRENT APPLICATION NUMBER: US/10/803,482
; CURRENT FILING DATE: 2004-03-18
; PRIOR FILING DATE: 2003-03-21
; NUMBER OF SEQ ID NOS: 231
; SEQ ID NO 28
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Antisense Oligonucleotide
US-10-803-482-28

Query Match          59.0%; Score 11.8; DB 18; Length 20;
Best Local Similarity 86.7%; Pred. No. 4e+04; 2; Indels 0; Gaps 0;
Matches 13; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GGGACGCGCGCTCG 15
Db 20 GGGTCGCGGCCCTCG 6

RESULT 12
US-10-803-482-103
; Sequence 103, Application US/10803482
; Publication No. US20040209838A1
; GENERAL INFORMATION:
; APPLICANT: Brett P. Monia
; APPLICANT: Mark J. Graham
; TITLE OF INVENTION: MODULATION OF DIACYLGLYCEROL ACYLTRANSFERASE 1 EXPRESSION
; FILE REFERENCE: BIOL000303
; CURRENT APPLICATION NUMBER: US/10/803,482
; CURRENT FILING DATE: 2004-03-18
; PRIOR FILING DATE: 2003-03-21
; NUMBER OF SEQ ID NOS: 231
; SEQ ID NO 103
; LENGTH: 20
; TYPE: RNA
; ORGANISM: H. sapiens
; FEATURE:
; OTHER INFORMATION: Antisense Oligonucleotide
US-10-803-482-103

Query Match          59.0%; Score 11.8; DB 18; Length 20;
Best Local Similarity 80.0%; Pred. No. 4e+04; 2; Indels 0; Gaps 0;
Matches 12; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GGGACGCGCGCTCG 15
Db 1 GGGUCGCGGCCCTCG 15
```


RESULT 15
US-09-942-596A-31
; Sequence 31, Application US/09942596A
; Patent No. US20020168648A1

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 3, 2005, 16:15:09 ; Search time 1802 Seconds
(without alignments)
537.794 Million cell updates/sec

Title: US-10-605-498-82

Perfect score: 20
Sequence: 1 gggacgcggcgctcggtcat 20

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 790860

Minimum DB seq length: 0
Maximum DB seq length: 20

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl.*
1: gb_ba.*
2: gb_htg.*
3: gb_in.*
4: gb_on.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_sy.*
13: gb_un.*
14: gb_vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	20	100.0	20	6	CQ799984	CQ799984 Sequence
c 2	19	95.0	19	6	CQ799992	CQ799992 Sequence
3	13	65.0	20	6	AR454276	AR454276 Sequence
4	12.4	62.0	20	6	AR014592	AR014592 Sequence
5	12.4	62.0	20	6	I26721	Sequence 9
6	12.4	62.0	20	6	AR235394	Sequence
7	12.4	62.0	20	6	BD010384	Chimeric
c 8	11.8	59.0	17	6	I34708	Sequence 6
9	11.8	59.0	19	6	AX202547	Sequence
c 10	11.8	59.0	20	6	AR370192	Sequence
11	11.6	58.0	18	6	BD133653	Method fo
12	11.6	58.0	18	6	BD135731	Method fo
13	11.6	58.0	18	6	BD160997	Terminal-
14	11.6	58.0	18	6	BD167492	A method
15	11.6	58.0	18	6	BD176975	Method of
c 16	11.6	58.0	18	6	AR488746	Sequence
17	11.6	58.0	18	6	AX391650	Sequence
18	11.6	58.0	18	6	AX391799	Sequence
c 19	11.6	58.0	18	6	AX419738	Sequence

20	11.6	58.0	18	6	AX453807	Sequence
21	11.6	58.0	18	6	BD000042	Probe-cou
c 22	11.6	58.0	18	6	BD085000	Target-de
23	11.6	58.0	20	6	AR271204	Sequence
24	11.4	57.0	17	6	AR138777	Sequence
c 25	11.4	57.0	17	6	AR138801	Sequence
26	11.4	57.0	20	6	AX078038	Sequence
c 27	11.2	56.0	17	6	AX760424	Sequence
c 28	11.2	56.0	19	6	BD016694	Method fo
29	11	55.0	15	6	AR180323	Sequence
c 30	11	55.0	20	6	AR208819	Sequence
31	11	55.0	20	6	AX412154	Sequence
32	11	55.0	20	6	AX546229	Sequence
33	10.8	54.0	15	6	BD208669	Enzymatic
c 34	10.8	54.0	16	6	AR435746	Sequence
c 35	10.8	54.0	17	6	I34707	Sequence 5
36	10.8	54.0	17	6	AR567813	Sequence
37	10.8	54.0	17	6	AX348120	Sequence
c 38	10.8	54.0	18	6	AR098803	Sequence
c 39	10.8	54.0	20	6	AR037484	Sequence
40	10.8	54.0	20	6	AR076701	Sequence
41	10.8	54.0	20	6	AR153558	Sequence
42	10.8	54.0	20	6	BD262873	Antisense
43	10.8	54.0	20	6	BD272875	Method of
44	10.8	54.0	20	6	AR182758	Sequence
45	10.8	54.0	20	6	AR300838	Sequence

ALIGNMENTS

RESULT 1
CQ799984
LOCUS CQ799984 20 bp DNA linear PAT 28-APR-2004
DEFINITION Sequence 82 from Patent WO2004030660.
ACCESSION CQ799984
VERSION CQ799984.1 GI:46848931
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Gleave, M.E., Rocchi, P. and Signaevsky, M.
TITLE Compositions for treatment of prostate and other cancers
JOURNAL Patent: WO 2004030660-A 82 15-APR-2004;
The University of British Columbia (CA)
FEATURES
source
1..20
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 100.0%; Score 20; DB 6; Length 20;
Best Local Similarity 100.0%; Pred. No. 8.4e+02;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GGGACGGCGCGCTCGGTCAAT 20
|||||
Db 1 GGGACGGCGCGCTCGGTCAAT 20
|||||

RESULT 2
CQ799992/c
LOCUS CQ799992 19 bp RNA linear PAT 28-APR-2004
DEFINITION Sequence 90 from Patent WO2004030660.
ACCESSION CQ799992
VERSION CQ799992.1 GI:46848939
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;


```
seeds of plants.
BD010384
VERSION BD010384.1 GI:18638757
KEYWORDS JP 2001502923-A/16.
SOURCE unidentified
ORGANISM unidentified
REFERENCE 1 (bases 1 to 20)
AUTHORS Falco,S.C., Iii,R.E.M. and Epelbaum,S.U.
TITLE Chimeric genes and methods for increasing the lysine content of the
JOURNAL seeds of plants
COMMENT Patent: JP 2001502923-A 16 06-MAR-2001;
EI DU PONT DE NEMOURS AND CO
OS Unidentified
PN JP 2001502923-A/16
PD 06-MAR-2001
PP 27-MAR-1998 JP 1998543284
PR 27-MAR-1997 US 08/824627
PI SAVERIO CARL FALCO,RAYMOND ERVIN MCDEVITT III, PI SABINE
URSLA EPELBAUM
PC C12N9/06,C12N9/12,C12N9/88,C12P13/08,C12N15/82 CC
Strandedness: Single;
CC Topology: Linear;
FH Key Location/Qualifiers
FT source 1..20
FT /organism="Unidentified".
FEATURES Location/Qualifiers
source 1..20
/organism="unidentified"
/mol_type="unassigned DNA"
/db_xref="taxon:32644"
ORIGIN
Query Match 62.0%; Score 12.4; DB 6; Length 20;
Best Local Similarity 92.9%; Pred. No. 1.2e+06;
Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 5 CGCGCGCTCGGTC 18
Db 7 CGTGGCTCGGTC 20
RESULT 8
I34708/c
LOCUS I34708 17 bp DNA linear PAT 13-MAY-1997
DEFINITION Sequence 6 from patent US 5599665.
ACCESSION I34708
VERSION I34708.1 GI:2087676
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 17)
AUTHORS Barbieri,J.T., Frank,D.W. and Kulich,S.M.
TITLE Pseudomonas aeruginosa nucleic acids encoding exoenzyme S activity
JOURNAL and use thereof in detecting pseudomonas aeruginosa infection
FEATURES Patent: US 5599665-A 6 04-FEB-1997;
source Location/Qualifiers
1..17
/organism="unknown"
/mol_type="unassigned DNA"
ORIGIN
Query Match 59.0%; Score 11.8; DB 6; Length 17;
Best Local Similarity 66.7%; Pred. No. 2.1e+06;
Matches 10; Conservative 4; Mismatches 1; Indels 0; Gaps 0;
Qy 2 GGACGCGCTCGGTC 16
Db 15 RGARGCGCTSGG 1
RESULT 9
I34708/c
LOCUS I34708 20 bp DNA linear PAT 12-SEP-2003
DEFINITION Sequence 13 from patent US 6300132.
ACCESSION AR370192
VERSION AR370192.1 GI:34606698
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 20)
AUTHORS Monia,B.P. and Cowsert,L.M.
TITLE Antisense inhibition of telomeric repeat binding factor 2
JOURNAL expression
FEATURES Patent: US 6300132-A 13 09-OCT-2001;
source Location/Qualifiers
1..20
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN
Query Match 59.0%; Score 11.8; DB 6; Length 20;
Best Local Similarity 86.7%; Pred. No. 2.1e+06;
Matches 13; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 5 CGCGCGCTCGGTCA 19
Db 18 CGCGCGATCGACA 4
RESULT 11
BD133653
LOCUS BD133653 18 bp DNA linear PAT 18-SEP-2002
DEFINITION Method for screening mutated gene.
ACCESSION BD133653
VERSION BD133653.1 GI:23228598
KEYWORDS JP 2002071687-A/31.
SOURCE synthetic construct
ORGANISM synthetic construct
REFERENCE 1 (bases 1 to 18)
other sequences; artificial sequences.
AX202547
LOCUS AX202547 19 bp DNA linear PAT 30-AUG-2001
DEFINITION Sequence 20 from Patent WO0153480.
ACCESSION AX202547
VERSION AX202547.1 GI:15392244
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE 1
AUTHORS Choo,Y., Klug,A. and Moore,M.T.
TITLE Nucleic acid binding polypeptides characterized by flexible linkers
JOURNAL connected nucleic acid binding modules
COMMENT Patent: WO 0153480-A 20 26-JUL-2001;
Gendaq Limited (GB)
FEATURES Location/Qualifiers
source 1..19
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Binding site sequence"
ORIGIN
Query Match 59.0%; Score 11.8; DB 6; Length 19;
Best Local Similarity 86.7%; Pred. No. 2.1e+06;
Matches 13; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 2 GGACGCGCGCTCGG 16
Db 3 GGACGCGCGCTTGG 17
RESULT 10
AR370192/c
LOCUS AR370192 20 bp DNA linear PAT 12-SEP-2003
DEFINITION Sequence 13 from patent US 6300132.
ACCESSION AR370192
VERSION AR370192.1 GI:34606698
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 20)
AUTHORS Monia,B.P. and Cowsert,L.M.
TITLE Antisense inhibition of telomeric repeat binding factor 2
JOURNAL expression
FEATURES Patent: US 6300132-A 13 09-OCT-2001;
source Location/Qualifiers
1..20
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN
Query Match 59.0%; Score 11.8; DB 6; Length 20;
Best Local Similarity 86.7%; Pred. No. 2.1e+06;
Matches 13; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 5 CGCGCGCTCGGTCA 19
Db 18 CGCGCGATCGACA 4
RESULT 11
BD133653
LOCUS BD133653 18 bp DNA linear PAT 18-SEP-2002
DEFINITION Method for screening mutated gene.
ACCESSION BD133653
VERSION BD133653.1 GI:23228598
KEYWORDS JP 2002071687-A/31.
SOURCE synthetic construct
ORGANISM synthetic construct
REFERENCE 1 (bases 1 to 18)
other sequences; artificial sequences.
```

```

AUTHORS Yamamoto,N., Okamoto,T., Suzuki,T. and Tanaka,S.
TITLE Method for screening mutated gene
JOURNAL Patent: JP 2002071687-A 31 12-MAR-2002;
COMMENT CANON INC
OS Artificial Sequence
PN JP 2002071687-A/31
PD 12-MAR-2002
PF 31-AUG-2000 JP 200263396
PI NOBUKO YAMAMOTO,TADASHI OKAMOTO,TOMOHIRO SUZUKI,SHINYA TANAKA
PC G01N33/53,C12M1/00,C12N15/09,C12Q1/68,G01N31/22,G01N33/566, PC
G01N37/00,
PC C12N15/00
CC Sample origonucleotide
FH Key Location/Qualifiers
FT source 1..18
FT /organism='Artificial Sequence'.
FEATURES
source
1..18
/organism='synthetic construct'
/mol_type='genomic DNA'
/db_xref='taxon:32630'
ORIGIN
Query Match 58.0%; Score 11.6; DB 6; Length 18;
Best Local Similarity 77.8%; Pred. No. 2.6e+06;
Matches 14; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 3 GACGGGGCTCGGTTCAT 20
DB 1 GATGGGGCTCTCGTTTCAT 18
RESULT 12
BD135731
LOCUS 18 bp DNA linear PAT 18-SEP-2002
DEFINITION Method for detecting subjective component in specimen sample, and
ACCESSION BD135731
VERSION JP 2002065274-A/35.
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE 1 (bases 1 to 18)
AUTHORS Yamamoto,N., Okamoto,T., Suzuki,T. and Shimizu,A.
TITLE Method for detecting subjective component in specimen sample, and
JOURNAL Substrate for detection used therefor
PATENT: JP 2002065274-A 35 05-MAR-2002;
COMMENT CANON INC
OS Artificial Sequence
PN JP 2002065274-A/35
PD 05-MAR-2002
PF 31-AUG-2000 JP 200263395
PI NOBUKO YAMAMOTO,TADASHI OKAMOTO,TOMOHIRO SUZUKI,AKIRA SHIMIZU
PC C12N15/09,C12M1/00,C12Q1/68,G01N31/22,G01N33/53, PC
G01N33/566,
PC G01N35/02,G01N35/10,G01N37/00,C12N15/00,G01N35/06 CC DNA
probe for hybridizing with gene encoding
mutated p53;named
CC in Table 1
CC as probe 31
FH Key Location/Qualifiers
FT source 1..18
FT /organism='Artificial Sequence'.
FEATURES
source
1..18
/organism='synthetic construct'
/mol_type='genomic DNA'
/db_xref='taxon:32630'
ORIGIN
Query Match 58.0%; Score 11.6; DB 6; Length 18;
Best Local Similarity 77.8%; Pred. No. 2.6e+06;
Matches 14; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 3 GACGGGGCTCGGTTCAT 20
DB 1 GATGGGGCTCTCGTTTCAT 18
RESULT 13
BD160997
LOCUS 18 bp DNA linear PAT 17-JAN-2003
DEFINITION Terminal-labeled probe-array and method for preparing it, and
method for evaluating target mass using the same.
ACCESSION BD160997
VERSION BD160997.1 GI:27866755
KEYWORDS JP 2002153284-A/31.
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE 1 (bases 1 to 18)
AUTHORS Okamoto,T., Yamamoto,N. and Suzuki,T.
TITLE Terminal-labeled probe-array and method for preparing it, and
JOURNAL method for evaluating target mass using the same
PATENT: JP 2002153284-A 31 28-MAY-2002;
COMMENT CANON INC
OS Artificial Sequence
PN JP 2002153284-A/31
PD 28-MAY-2002
PF 24-NOV-2000 JP 2000357446
PI TADASHI OKAMOTO,NOBUKO YAMAMOTO,TOMOHIRO SUZUKI PC
C12N15/09,C12Q1/68,G01N31/22,G01N33/53,G01N37/00, PC
C12N15/00
CC Description of Artificial Sequence:Synthesized FH Key
Location/Qualifiers
FT source 1..18
FT /organism='Artificial Sequence'.
FEATURES
source
1..18
/organism='synthetic construct'
/mol_type='genomic DNA'
/db_xref='taxon:32630'
ORIGIN
Query Match 58.0%; Score 11.6; DB 6; Length 18;
Best Local Similarity 77.8%; Pred. No. 2.6e+06;
Matches 14; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 3 GACGGGGCTCGGTTCAT 20
DB 1 GATGGGGCTCTCGTTTCAT 18
RESULT 14
BD167492
LOCUS 18 bp DNA linear PAT 17-JAN-2003
DEFINITION A method of analyzing a base sequence of a nucleic acid.
ACCESSION BD167492
VERSION BD167492.1 GI:27873304
KEYWORDS WO 0233068-A/31.
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE 1 (bases 1 to 18)
AUTHORS Yamamoto,N., Okamoto,T. and Suzuki,T.
TITLE A method of analyzing a base sequence of a nucleic acid
JOURNAL Patent: WO 0233068-A 31 25-APR-2002;
COMMENT CANON KK,NOBUKO YAMAMOTO,TADASHI OKAMOTO,TOMOHIRO SUZUKI
OS Artificial Sequence
PN WO 0233068-A/31
PD 25-APR-2002
PF 18-OCT-2000 WO 2000JP007244
PI NOBUKO YAMAMOTO,TADASHI OKAMOTO,TOMOHIRO SUZUKI PC
C12N15/09,C12Q1/68,G01N33/566,G01N33/53
CC Sample origonucleotide

```

Search completed: May 3, 2005, 20:18:53
Job time : 1810 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 3, 2005, 16:10:24 ; Search time 422 Seconds
(without alignments)
280.556 Million cell updates/sec

Title: US-10-605-498-82

Perfect score: 20

Sequence: 1 gggacgcgcgcgcgtcat 20

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 2207178

Minimum DB seq length: 0

Maximum DB seq length: 20

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_16Dec04.*

- 1: Geneseqn1980s.*
- 2: Geneseqn1990s.*
- 3: Geneseqn2000s.*
- 4: Geneseqn2001as.*
- 5: Geneseqn2001bs.*
- 6: Geneseqn2002as.*
- 7: Geneseqn2002bs.*
- 8: Geneseqn2003as.*
- 9: Geneseqn2003bs.*
- 10: Geneseqn2003cs.*
- 11: Geneseqn2003ds.*
- 12: Geneseqn2004as.*
- 13: Geneseqn2004bs.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	20	100.0	20	12 ADM94732	Adm94732 Human hea
c 2	19	95.0	19	12 ADM94740	Adm94740 Human hea
3	12.6	63.0	20	6 ABZ08781	Abz08781 Human CMV
c 4	12.4	62.0	19	3 AAA40594	Aaa40594 Human Arp
5	12.4	62.0	20	2 AAQ36810	Aaq36810 Oligomer
6	12.4	62.0	20	2 AAQ94968	Aaq94968 PCR prime
7	12.4	62.0	20	2 AAV35805	AAv35805 PCR prime
8	12.4	62.0	20	2 AAV95501	AAv95501 Plasmid p
9	12.4	62.0	20	8 ADA16031	Ada16031 Plasmid p
10	12.4	62.0	20	9 ACH03677	Ach03677 Plasmid p
c 11	12.4	62.0	20	10 ADF73001	Adf73001 Probe rel
c 12	11.8	59.0	15	6 AAD26768	Aad26768 Human SNA
c 13	11.8	59.0	17	2 AAT61840	Aat61840 Pseudomon
14	11.8	59.0	19	4 AAH23296	Aah23296 Coding st
c 15	11.8	59.0	20	5 AAH20644	Aah20644 Human tel
16	11.8	59.0	20	10 ABZ87221	Abz87221 Human oli
17	11.8	59.0	20	11 ABD23451	Abd23451 Human myo
c 18	11.8	59.0	20	13 ADS00264	Ads00264 Human dia
19	11.8	59.0	20	13 ADS00339	Ads00339 Human dia
c 20	11.6	58.0	18	2 AAV70497	AAv70497 Truncated

c	21	11.6	58.0	18	3	AAAS8530	Aaa858530 PCR prime
	22	11.6	58.0	18	4	AAC99272	Aac99272 Probe seq
	23	11.6	58.0	18	6	ABK72469	Abk72469 Sample or
	24	11.6	58.0	18	6	ABN99777	Abn99777 DNA probe
	25	11.6	58.0	18	6	ABL54931	AbL54931 Human tum
c	26	11.6	58.0	18	6	ABL46108	AbL46108 Mycobacte
	27	11.6	58.0	18	6	ABT04724	Abt04724 End-label
	28	11.6	58.0	18	6	ABL59666	AbL59666 Oligonuc1
	29	11.6	58.0	18	6	ABT06245	Abt06245 Synthetic
c	30	11.6	58.0	18	12	ADK82298	Adk82298 Mycobacte
c	31	11.6	58.0	20	4	AAH22265	Aah22265 Fibroblae
	32	11.6	58.0	20	10	ADC65851	Adc65851 Mouse TGF
c	33	11.6	58.0	20	10	ABZ98678	Abz98678 Human try
	34	11.6	58.0	20	10	ABZ87769	Abz87769 Human oli
c	35	11.6	58.0	20	11	ABD31709	Abd31709 Human Try
	36	11.6	58.0	20	11	ABD23999	Abd23999 Human cal
c	37	11.6	58.0	20	12	ADJ60557	Adj60557 Oligonuc1
c	38	11.6	58.0	20	12	ADO46046	Ado46046 Human oli
	39	11.4	57.0	13	12	ADM82136	Adm82136 Monofiliam
	40	11.4	57.0	13	12	ADQ76718	Adq76718 Enzymatic
c	41	11.4	57.0	15	6	ABK95806	Abk95806 Solute Ca
	42	11.4	57.0	15	6	ABI99079	Abi99079 Human PCD
c	43	11.4	57.0	17	2	AAV09529	AAv09529 MSP ampli
	44	11.4	57.0	17	2	AAV09425	AAv09425 Cpg-conta
c	45	11.4	57.0	17	12	ADQ78137	Adq78137 PCR prime

ALIGNMENTS

RESULT 1

ADM94732
ID ADM94732 standard; DNA; 20 BP.
XX
AC ADM94732;
XX
DT 01-JUL-2004 (first entry)
XX
DE Human heat shock protein 27 antisense oligonucleotide SEQ ID NO:82.
XX
KW heat shock protein 27; hsp27; cytosolic; gene therapy;
KW heat shock protein 27 inhibitor; hsp27 inhibitor; cancer; human;
KW antisense oligonucleotide; ss.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN WO2004030660-A2.
XX
PD 15-APR-2004.
XX
PF 02-OCT-2003; 2003WO-CA001588.
XX
PR 02-OCT-2002; 2002US-0415859P.
PR 18-APR-2003; 2003US-0463952P.
XX
PA (UYBR-) UNIV BRITISH COLUMBIA.
XX
PI Gleave ME, Rocchi P, Signaevsky M;
XX
DR WPI; 2004-316331/29.
XX
PT New composition comprising a therapeutic agent that reduces the amount of active hsp27 in hsp27 expressing cells exposed to the therapeutic agent, useful in treating cancer, e.g., prostate cancer or a central nervous system malignancy.
XX
PS Claim 6; SEQ ID NO 82; 38pp; English.
XX
CC The present invention describes a composition which comprises a therapeutic agent that reduces the amount of active heat shock protein 27 (hsp27) in hsp27 expressing cells exposed to the therapeutic agent. The composition has cytostatic activity, and can be used in gene therapy. The

CC composition is useful in treating cancer, e.g., prostate, bladder, lung, breast, pancreatic, colon, skin (for example melanoma), renal or ovarian cancer or a central nervous system malignancy. The present sequence CC represents a human hsp27 antisense oligonucleotide which is used in the CC exemplification of the present invention.

XX Sequence 20 BP; 2 A; 6 C; 9 G; 3 T; 0 U; 0 Other;

Query Match 100.0%; Score 20; DB 12; Length 20;

Best Local Similarity 100.0%; Pred. No. 11;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGACGGCGCTCGGTTCAT 20

DB 1 GGGACGGCGCTCGGTTCAT 20

RESULT 2

ADM94740/c

ID ADM94740 standard; DNA; 19 BP.

XX

AC ADM94740;

DT 01-JUL-2004 (first entry)

XX Human heat shock protein 27 siRNA oligonucleotide SEQ ID NO:90.

DE heat shock protein 27; hsp27; cytostatic; gene therapy;

XX heat shock protein 27 inhibitor; hsp27 inhibitor; cancer; human;

KW short interfering RNA; siRNA; RNA interference; RNAi; ds.

XX Homo sapiens.

OS Synthetic.

XX WO2004030660-A2.

PN 15-APR-2004.

XX 02-OCT-2003; 2003WO-CA001588.

PF 02-OCT-2002; 2002US-0415859P.

XX 18-APR-2003; 2003US-0463952P.

PR (UYBR-) UNIV BRITISH COLUMBIA.

XX Gleave ME, Rocchi P, Signaevsky M;

PI WPI; 2004-316331/29.

XX New composition comprising a therapeutic agent that reduces the amount of

DR active hsp27 in hsp27 expressing cells exposed to the therapeutic agent,

XX useful in treating cancer, e.g., prostate cancer or a central nervous

PT system malignancy.

XX Claim 10; SEQ ID NO 90; 38pp; English.

PS The present invention describes a composition which comprises a

XX therapeutic agent that reduces the amount of active heat shock protein 27

CC (hsp27) in hsp27 expressing cells exposed to the therapeutic agent. The

CC composition has cytostatic activity, and can be used in gene therapy. The

CC composition is useful in treating cancer, e.g., prostate, bladder, lung,

CC breast, pancreatic, colon, skin (for example melanoma), renal or ovarian

CC cancer or a central nervous system malignancy. The present sequence

CC represents a human hsp27 short interfering RNA (siRNA) oligonucleotide

CC which is used in the exemplification of the present invention.

XX Sequence 19 BP; 3 A; 8 C; 6 G; 0 T; 2 U; 0 Other;

Query Match 95.0%; Score 19; DB 12; Length 19;

Best Local Similarity 100.0%; Pred. No. 30;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX 2 GGGACGGCGCTCGGTTCAT 20

QY

DB 19 GGGACGGCGCTCGGTTCAT 1

RESULT 3

ABZ08781

ID ABZ08781 standard; DNA; 20 BP.

XX

AC ABZ08781;

DT 09-JAN-2003 (first entry)

XX Human CMV PCR primer SEQ ID NO 8773.

DE CMV; leukocyte; gene expression profiling; allograft rejection;

XX atherosclerosis; congestive heart failure; systemic lupus erythematosus;

KW rheumatoid arthritis; osteoarthritis; cytomegalovirus; infection; PCR;

XX primer; ss.

XX Human cytomegalovirus.

OS WO200257414-A2.

PN 25-JUL-2002.

XX 22-OCT-2001; 2001WO-US047856.

PF 20-OCT-2000; 2000US-0241994P.

XX 08-JUN-2001; 2001US-0296764P.

PR (BIOC-) BIOCARDIA INC.

XX Wohlgemuth J, Fry K, Matcuk G, Altman P, Prentice J, Phillips J;

PI Ly N, Woodward R, Quattermost T, Johnson F;

XX WPI; 2002-636525/68.

DR New system for leukocyte expression profiling, diagnosing a disease, or

XX monitoring (the rate of) progression of a disease, e.g. atherosclerosis

PT or congestive heart failure, comprises diagnostic oligonucleotides.

XX Example 18; Page 142; Opp; English.

PS The invention relates to a system for detecting gene expression, which

XX comprises one or two isolated DNA molecules that detect expression of a

CC gene, where the gene corresponds to any of 8143 oligonucleotides

CC (ABZ0010-ABZ08152) each having 50 base pairs (bp). The system is useful

CC for leukocyte expression profiling. It is particularly useful for

CC diagnosing a disease, monitoring (rate of) progression of a disease,

CC predicting therapeutic outcome, determining prognosis for a patient,

CC predicting disease complications in an individual or monitoring response

CC to treatment in an individual. The diseases include cardiac allograft

CC rejection, kidney allograft rejection, liver allograft rejection,

CC atherosclerosis, congestive heart failure, systemic lupus erythematosus,

CC rheumatoid arthritis, osteoarthritis or cytomegalovirus infection. The

CC present sequence is that of a CMV PCR primer used in the invention

XX Sequence 20 BP; 2 A; 5 C; 8 G; 5 T; 0 U; 0 Other;

QY Query Match 63.0%; Score 12.6; DB 6; Length 20;

Best Local Similarity 78.9%; Pred. No. 2.7e+04;

Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

DB 2 GGGACGGCGCTCGGTTCAT 20

1 GTACGGCTCGGTTCAT 19

QY

RESULT 4

AAA40594/c

ID AAA40594 standard; DNA; 19 BP.

XX

AC AAA40594;

XX 20-NOV-2000 (first entry)
 XX Human Arp3H PCR primer #1.
 XX Human; Arp3H; actin-related protein; PCR primer; amplification; ss.
 XX Homo sapiens.
 XX CN1252448-A.
 XX 10-MAY-2000.
 XX 22-OCT-1998; 98CN-00123482.
 XX 22-OCT-1998; 98CN-00123482.
 XX (XINH-) XINHUANGFU FUDAN GENE ENG CO LTD SHANGHA.
 XX Yu L, Fu Q, Zhang H;
 XX WPI; 2000-443229/39.
 XX Human actin related protein gene and encoded polypeptide preparation.
 XX Example 1; Page 17; 25pp; Chinese.
 XX This invention describes a novel preparation which provides a new human
 CC Arp3H as an actin-related protein gene family member. The present
 CC invention provides the cDNA sequence of the actin related protein gene
 CC and the polypeptide encoded by the said sequence, as well as the
 CC recombination technology to produce the actin related protein gene. This
 CC sequence represents a PCR primer used in the amplification of the human
 CC Arp3A gene described in the method of the invention
 XX
 SQ Sequence 19 BP; 1 A; 8 C; 10 G; 0 T; 0 U; 0 Other;
 Query Match 62.0%; Score 12.4; DB 3; Length 19;
 Best Local Similarity 92.9%; Pred. No. 3.3e+04;
 Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 5 CGCGGCGCTCGGTC 18
 |||||
 Db 17 CGCGGCGCTCGGCC 4
 RESULT 5
 ID AAQ36810 standard; DNA; 20 BP.
 XX AAQ36810;
 AC AAQ36810;
 XX 25-MAR-2003 (revised)
 DT 22-JUN-1993 (first entry)
 XX Oligomer SM 70 used in construction of SSP polypeptides.
 DE Heptad; plants; custom tailored storage proteins; in vivo; expression;
 KW ss.
 XX Synthetic.
 OS W09303160-A1.
 PN 18-FEB-1993.
 PD 07-AUG-1992; 92WO-US006412.
 XX 09-AUG-1991; 91US-00743006.
 PR (DUPO) DU PONT DE NEMOURS & CO E I.
 XX Falco SC, Keeler SJ, Rice JA;
 PI

XX WPI; 1993-076517/09.
 DR Synthetic polypeptide(s) contg. specified heptad units - expressed in
 XX vivo in plants to serve as custom-tailored storage proteins with
 PT specified aminoacid content.
 XX Disclosure; Page 106; 176pp; English.
 PS The sequence represents the DNA sequence encoding a synthetic heptad
 CC polypeptide. The synthetic polypeptide can be expressed in vivo in plants
 CC to serve as a synthetic seed storage protein which can be custom-tailored
 CC for specific end-user requirements. The DNA encoding the heptad may be
 CC used to transform plants to increase the content of partic. amino acids
 CC such as lysine or methionine in seeds or leaves. See also AAQ36811-28,
 CC AAQ37265-301. (Updated on 25-MAR-2003 to correct PN field.)
 XX
 SQ Sequence 20 BP; 1 A; 8 C; 6 G; 5 T; 0 U; 0 Other;
 Query Match 62.0%; Score 12.4; DB 2; Length 20;
 Best Local Similarity 92.9%; Pred. No. 3.3e+04;
 Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 5 CGCGGCGCTCGGTC 18
 |||||
 Db 7 CGCTGCGCTCGGTC 20
 RESULT 6
 ID AAQ94968 standard; DNA; 20 BP.
 XX AAQ94968;
 AC AAQ94968;
 XX 15-JUL-1996 (first entry)
 DT PCR primer SM 70.
 DE Lysine; synthetic storage protein; SSP; vector; pSK6;
 KW dihydridipicolinic acid synthase; corn; maize; Zea mays; soybean;
 KW Glycine max; rapeseed; oilseed rape; Brassica napus; transgenic plant;
 KW essential amino acid; primer; PCR; polymerase chain reaction; ss.
 XX Synthetic.
 OS
 XX Key Location/Qualifiers
 FH misc_feature 1..20
 FT /*tag= a
 FT /standard_name= "SM 70"
 XX W09515392-A1.
 PN 08-JUN-1995.
 PD 21-NOV-1994; 94WO-US013190.
 XX 30-NOV-1993; 93US-00160117.
 PR 17-JUN-1994; 94US-00261661.
 XX (DUPO) DU PONT DE NEMOURS & CO E I.
 PA Falco SC, Keeler SJ, Rice JA;
 XX WPI; 1995-215272/28.
 DR New chimeric gene providing increased lysine content in plant seeds -
 PT contains di:hydro:di:picolinic acid synthase gene coupled to chloroplast
 PT transport sequence and seed specific promoter, also new plants of
 PT improved nutritional value.
 XX Example 8; Page 71; 180pp; English.
 PS 2 PCR primers (AAQ94968 and AAQ94969) were designed to prime replication
 XX

CC of plasmid pSK1, but excluding a 30 base fragment around the EarI site.
 CC pSK1 is a mutant of pBR322 in which the ampicillin gene has been deleted.
 CC PCR was performed using pSK1 as template. The product was utilised in the
 CC prodn. of pSK6 (see also AAQ94970-71), a vector used for the construction
 CC of genes (see AAQ94972-Q95005) coding for synthetic storage proteins
 CC (SSPs) (see AAR78236-59). These were expressed in the seeds of
 CC transformed soybean, rapeseed and corn to improve the lysine content
 XX
 SQ Sequence 20 BP; 1 A; 8 C; 6 G; 5 T; 0 U; 0 Other;

Query Match 62.0%; Score 12.4; DB 2; Length 20;
 Best Local Similarity 92.9%; Pred. No. 3.3e+04;
 Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 CGCGCGCTCGGTC 18
 ||| |||||
 Db 7 CGCTCGCTCGGTC 20

RESULT 7
 AAV35805
 ID AAV35805 standard; DNA; 20 BP.
 XX AC AAV35805;
 XX 25-MAR-2003 (revised)
 DT 25-AUG-1998 (first entry)
 XX PCR primer used in the course of the invention.
 XX Dihydrodipicolinic acid synthase; DHDPS; chimeric gene;
 KW lysine inhibition; plant chloroplast transit sequence;
 KW plant seed-specific regulatory sequence; transgenic plant;
 KW increased lysine level; corn; Zea mays; soybean; Glycine max; PCR primer;
 KW ss.

XX Synthetic.
 OS US5773691-A.
 PN 30-JUN-1998.

XX 07-JUN-1995; 95US-00474633.
 XX 19-MAR-1992; 92US-00855414.
 PR 18-MAR-1993; 93WO-US002480.
 PR 06-JAN-1994; 94US-00178212.

XX (DUPO) DU PONT DE NEMOURS & CO E I.

XX Falco SC;
 PI WPI; 1998-387117/33.

XX Chimeric genes encoding lysine production enzymes - useful for increasing
 PT transgenic seed lysine content without being inhibited by high levels of
 PT the amino acid.

XX Example 21; Col 73; 106pp; English.

XX PCR primers AAV35805-06 were used to in the course of the invention. The
 CC specification describes a Corynebacterium dapsa gene, which encodes a
 CC dihydrodipicolinic acid synthase (DHDPS) enzyme, which was used to create
 CC chimeric genes of the invention. The chimeric genes contain a nucleic
 CC acid fragment encoding a DHDPS enzyme which is insensitive to inhibition
 CC by lysine operably linked to a plant chloroplast transit sequence and to
 CC a plant seed-specific regulatory sequence. The chimeric genes are useful
 CC for producing plants containing increased levels of lysine, especially in
 CC corn (Zea mays) and soybean (Glycine max). (Updated on 25-MAR-2003 to
 CC correct PR field.)

XX Sequence 20 BP; 1 A; 8 C; 6 G; 5 T; 0 U; 0 Other;

Query Match 62.0%; Score 12.4; DB 2; Length 20;
 Best Local Similarity 92.9%; Pred. No. 3.3e+04;
 Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 CGCGCGCTCGGTC 18
 ||| |||||
 Db 7 CGCTCGCTCGGTC 20

RESULT 8
 AAV99501
 ID AAV99501 standard; DNA; 20 BP.

XX AC AAV99501;

XX 29-MAR-1999 (first entry)

XX Plasmid pSK1 PCR primer SM70.

XX Lysine; transgenic plant; seed storage protein; vector; pSK5; PCR;
 KW primer; ss.

XX Synthetic.
 OS Escherichia coli.

XX WO9842831-A2.

XX 01-OCT-1998.

XX 27-MAR-1998; 98WO-US006051.

XX 27-MAR-1997; 97US-00824627.

XX (DUPO) DU PONT DE NEMOURS & CO E I.

XX Falco SC, Mcdevitt RE, Epelbaum SU;

XX WPI; 1999-045139/04.

XX Nucleic acids and chimeric genes for increasing seed lysine content -
 PT comprise sequence encoding all or part of lysine ketoglutarate reductase,
 PT useful to improve nutritional quality of seeds from transformed plants.

XX Example 21; Page 99; 231pp; English.

XX PCR primers SM70 and SM71 (see AAV99502) are designed to remove the EarI
 CC site at base 2353 of plasmid pSK1, a spontaneous mutant of pBR322 in
 CC which the ampicillin gene and the EarI site near the gene are deleted.
 CC The product was used in the construction of expression vector pSK5.

XX Chimeric genes for lysine-rich synthetic seed storage proteins suitable
 CC for expression in the seeds of plants have been constructed in pSK5 (see
 CC AAV99513-18, AAV99527-32, AAV99539-41). The invention also provides
 CC claimed nucleic acids and chimeric genes useful for improving the
 CC nutritional quality of seeds from transgenic plants. Methods involve
 CC manipulation of lysine ketoglutarate reductase and lysine-insensitive
 CC dihydrodipicolinic acid synthase activity (see AAW87757-66)

XX Sequence 20 BP; 1 A; 8 C; 6 G; 5 T; 0 U; 0 Other;

Query Match 62.0%; Score 12.4; DB 2; Length 20;
 Best Local Similarity 92.9%; Pred. No. 3.3e+04;
 Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 CGCGCGCTCGGTC 18
 ||| |||||
 Db 7 CGCTCGCTCGGTC 20

RESULT 9
 ADA16031
 ID ADA16031 standard; DNA; 20 BP.

XX ADA16031;

```

XX 06-NOV-2003 (first entry)
XX Plasmid pSK1 mutagenic PCR primer #1.
XX ss; lycC; transgenic; lysine accumulation;
XX dihydrodipicolinic acid synthase; DHDPs; lysine inhibition;
XX lysine ketoglutarate reductase; LKR; chloroplast transit sequence; CTS;
XX aspartokinase III; AKIII; synthetic seed storage protein; SSP.
XX Synthetic.
XX US6459019-B1.
XX 01-OCT-2002.
XX 24-MAR-1997; 97US-00823771.
XX 19-MAR-1992; 92US-00855414.
XX 06-JAN-1994; 94US-00178212.
XX 07-JUN-1995; 95US-00474633.
XX (DUPO ) DU PONT DE NEMOURS & CO E. I.
XX Falco SC, Keeler SJ, Rice JA;
XX WPI; 2003-028272/02.
XX Transformed plants that accumulate lysine at higher levels in its seeds
XX than untransformed plants, has gene fragments encoding lysine-insensitive
XX dihydrodipicolinic acid synthase and lysine ketoglutarate reductase.
XX Example 21; Col 76; 109pp; English.
XX The invention relates to a plant comprising two foreign nucleotide
XX sequences which cause seeds obtained from the plant to accumulate lysine
XX at a level of at least 10% higher than seeds of a plant that do not
XX comprise the nucleotide, where the nucleotide comprises a fragment
XX encoding a dihydrodipicolinic acid synthase (DHDPs) that is insensitive
XX to lysine inhibition, and a fragment encoding a plant lysine
XX ketoglutarate reductase (LKR) or its subfragment. The nucleotide fragment
XX is operably linked to a plant chloroplast transit sequence (CTS) and the
XX plant lysine ketoglutarate reductase subfragment is used in antisense
XX inhibition or cosuppression. Also included are progeny plants from the
XX above mentioned plant and seeds obtained from the above mentioned plant.
XX The seeds obtained from the above mentioned plant (e.g., rapeseed,
XX soybean or corn) comprising the foreign nucleic acid sequences accumulate
XX lysine at a higher level, preferably at a level of at least 10% higher
XX than seeds of a plant that do not comprise the foreign nucleic acid
XX sequences. Chimaeric gene comprising DHDPs from C. glutamicum and
XX aspartokinase III (from the lycC gene) of E. coli (mutated to be lysine-
XX insensitive) are also used to generate the above transgenic plants. Also
XX disclosed are synthetic seed storage proteins (SSP) used as an internal
XX source of lysine, built up from synthetic peptide monomers based around
XX an EarI site sequence (for generating multimeric proteins). The present
XX sequence is an oligonucleotide (or primer) used in the construction of
XX expression plasmids for expression of the chimaeric proteins or SSPs.
XX Sequence 20 BP; 1 A; 8 C; 6 G; 5 T; 0 U; 0 Other;
XX
XX Query Match 62.0%; Score 12.4; DB 8; Length 20;
XX Best Local Similarity 92.9%; Pred. No. 3.3e+04;
XX Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
XX QY 5 CCGCGCGCTCGGTC 18
XX ||| |||||
XX Db 7 CCGCTCGGCTCGGTC 20
XX
XX RESULT 10
XX ACH03677
XX ID ACH03677 standard; DNA; 20 BP.
XX

```

```

AC ACH03677;
XX 25-SEP-2003 (first entry)
XX Plasmid pSK1 Ear I site mutagenic PCR primer #1.
XX Aspartokinase; AKIII; dihydrodipicolinic acid synthase; DHDPs;
XX seed lysine content; seed threonine content; seed storage protein; SSP;
XX chloroplast transit sequence; lysine-rich protein;
XX lysine ketoglutarate reductase; LKR; transgenic; ss; PCR; primer.
XX Escherichia coli.
XX Synthetic.
XX US2003056242-A1.
XX 20-MAR-2003.
XX 17-DEC-2001; 2001US-00023066.
XX 19-MAR-1992; 92US-00855414.
XX 18-MAR-1993; 93WO-US002480.
XX 06-JAN-1994; 94US-00178212.
XX 07-JUN-1995; 95US-00474633.
XX 24-MAR-1997; 97US-00823771.
XX (FALC/) FALCO S C.
XX Falco SC;
XX WPI; 2003-521869/49.
XX New nucleic acid fragment encoding aspartokinase and dihydrodipicolinic
XX acid synthase, useful for increasing threonine or lysine content of seeds
XX of plant.
XX Example 21; Page 41; 116pp; English.
XX The invention relates to an isolated nucleic acid fragment comprising a
XX first nucleic acid subfragment encoding aspartokinase (AK) that is
XX substantially insensitive to inhibition by lysine, and a second nucleic
XX acid subfragment encoding dihydrodipicolinic acid synthase (DHDPs) that
XX is substantially insensitive to inhibition by lysine. Also included are
XX an isolated nucleic acid fragment comprising a nucleic acid subfragment
XX encoding lysine ketoglutarate reductase (LKR), a chimaeric gene (where
XX the nucleic acid fragment is operably linked to a plant chloroplast
XX transit sequence and to a seed-specific regulatory sequence, a plant
XX comprising the nucleic acid/chimaeric gene in its genome, a seed obtained
XX from the plant, increasing threonine or lysine content of the seeds of
XX plant, a plant capable of transmitting the chimaeric gene to a progeny of
XX plant having the ability to produce levels of free threonine or lysine at
XX least two times greater than the free threonine levels of untransformed
XX plants, a transformed (soybean) plant comprising seeds that accumulate
XX lysine at a level at least ten percent to four-fold higher than the seeds
XX of an untransformed plant, a transformed rapeseed comprising seeds that
XX accumulate lysine to a level between ten percent and one hundred percent
XX higher than that of the seeds of an untransformed plant, a monocot plant
XX comprising in its genome the nucleic acid fragment having the monocot-
XX embryo specific promoter and a transformed corn plant comprising seeds
XX that accumulate lysine to a level between ten percent and one hundred
XX thirty percent higher than the seeds of the untransformed plant. Also
XX disclosed are synthetic lysine-rich seed storage proteins (SSP), built up
XX from monomer lysine-rich heptad repeats (encoded by EarI restriction
XX enzyme-based oligonucleotides) used as a pool of lysine in a transformed
XX plant. The nucleic acid fragments, genes and methods are useful for
XX increasing threonine or lysine content of the seeds of the plant. Seeds
XX containing increased threonine or lysine content eliminate the need to
XX supplement mixed grain feeds with lysine or threonine produced via
XX microbial fermentation. The present sequence is an oligonucleotide used
XX in the preparation of a version of plasmid pSK1 used to clone the SSP
XX synthetic sequences
XX Sequence 20 BP; 1 A; 8 C; 6 G; 5 T; 0 U; 0 Other;
XX

```

```
Query Match      62.0%; Score 12.4; DB 9; Length 20;
Best Local Similarity 92.9%; Pred. No. 3.3e+04;
Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 CGCGGGCTCGGTC 18
DB 7 CGCTGGCTCGGTC 20

RESULT 11
ADF73001/c
ID ADF73001 standard; DNA; 20 BP.
XX
AC ADF73001;
XX
DT 26-FEB-2004 (first entry)
XX
DE Probe related to the invention #59.
XX
KW tubercle bacillus gene chip; ss; probe.
XX
OS Synthetic.
XX
PN CN1362526-A.
XX
PD 07-AUG-2002.
XX
PF 05-JAN-2001; 2001CN-00107010.
XX
PR 05-JAN-2001; 2001CN-00107010.
XX
PA (BAOL/) BAO L.
XX
PI Bao L, Zhang W, Wang X;
XX
DR WPI; 2003-240333/24.
XX
PT A tubercle bacillus gene chip useful for tuberculosis diagnosis and
PT reasonable selection of medicine.
XX
PS Disclosure; SEQ ID NO 59; 4pp; Chinese.
XX
CC The present sequence represents a tubercle bacillus gene chip and its
CC application. According to the characteristics of tubercle bacillus genome
CC sequence and molecular mechanism produced by tubercle bacillus resistance
CC to drug, and according to its application the probe can be designed and
CC selected, and the probe array can be regularly and reasonable arranged
CC according to a certain mode to form optimized probe array. The gene chip
CC can be used for quickly, accurately and high-efficiency identifying
CC tubercle bacillus, and can be used for screening and detecting its
CC resistance to drug, and can detect the mutation of specific site of
CC tubercle bacillus gene sequence and the mutation of non-specific site.
CC The invention can be used for tuberculosis diagnosis and reasonable
CC selection of medicine. The present sequence represents a probe related to
CC the invention.
XX
SQ Sequence 20 BP; 3 A; 10 C; 5 G; 2 T; 0 U; 0 Other;

Query Match      62.0%; Score 12.4; DB 10; Length 20;
Best Local Similarity 92.9%; Pred. No. 3.3e+04;
Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 6 CGCGGGCTCGGTC 19
DB 19 CGCGGGCTCGGTC 6

RESULT 12
AAD26768/c
ID AAD26768 standard; DNA; 15 BP.
XX
AC AAD26768;
XX
```

```
XX 09-APR-2002 (first entry)
XX Human SNAP29 gene polymorphism detecting ASO primer #11.
XX
KW Human; synaptosomal-associated protein 29 kD; SNAP29 protein; primer;
KW haplotyping; genotyping; membrane trafficking related disorder; ASO;
KW gene therapy; polymorphism; allele-specific oligonucleotide; ss.
XX
OS Homo sapiens.
XX
PN WO200190126-A2.
XX
PD 29-NOV-2001.
XX
PF 23-MAY-2001; 2001WO-US016938.
XX
PR 23-MAY-2000; 2000US-0206529P.
XX
PA (GENA-) GENAISSANCE PHARM INC.
XX
PI Kazemi A, Koshiy B, Tanguay DA;
XX
DR WPI; 2002-089916/12.
XX
CC Novel isolated polynucleotide, a polymorphic variant of synaptosomal-
CC associated protein, 29 kD (SNAP29) gene useful for expressing SNAP29
CC protein isoform to screen drugs to treat SNAP29 activity-related disease.
XX
PS Claim 16; Page 13; 115pp; English.
XX
CC The invention relates to genetic variants of human synaptosomal-
CC associated protein, 29 kD (SNAP29) gene. The invention also relates to
CC compositions and methods for haplotyping and/or genotyping the SNAP29
CC gene in an individual. Polynucleotides of the invention are useful for
CC studying the expression and function of SNAP29 and expressing SNAP29
CC protein for use in screening candidate drugs to treat diseases related to
CC SNAP29 activity. They are also used in gene therapy. The genotyping
CC method is useful for determining if an individual has one of haplotype or
CC haplotype pairs. The polymorphism and haplotype data is useful for
CC validating whether SNAP29 is a suitable target for drugs to treat
CC disorders related to membrane trafficking, screening for such drugs and
CC reducing bias cells in clinical trials of such drugs. The present
CC sequence is an allele-specific oligonucleotide (ASO) primer used to
CC detect human SNAP29 gene polymorphisms
XX
SQ Sequence 15 BP; 3 A; 6 C; 5 G; 0 T; 0 U; 1 Other;

Query Match      59.0%; Score 11.8; DB 6; Length 15;
Best Local Similarity 86.7%; Pred. No. 6.3e+04;
Matches 13; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 GACGGGGCTCGGCT 17
DB 15 GSCGGGGCTCGGCT 1

RESULT 13
AAT61840/c
ID AAT61840 standard; DNA; 17 BP.
XX
AC AAT61840;
XX
DT 08-MAY-1997 (first entry)
XX
DE Pseudomonas aeruginosa exoenzyme S gene antisense PCR primer.
XX
KW Exoenzyme S; exoS; detection; treatment; vaccination; infection;
KW production; antibody; modification; ras protein; carcinoma; probe;
KW mammalian; recombinant; primer; PCR; polymerase chain reaction; ss.
XX
OS Synthetic.
XX
```

```
PN US5599665-A.
XX
PD 04-FEB-1997.
XX
PF 21-DEC-1993; 93US-00171299.
XX
PR 21-DEC-1993; 93US-00171299.
XX
PA (MCWR-) MCW RES FOUND INC.
XX
PI Kulich SM, Barbieri JT, Frank DW;
XX
DR WPI; 1997-118290/11.
XX
XX New isolated P. aeruginosa exo:enzyme S DNA - used to develop prods. for
PT the detection, treatment and prevention of infections or for treating
PT carcinoma(s).
XX
XX Disclosure; Col 19-20; 23pp; English.
XX
CC The present sequence is a PCR primer for the gene encoding the 49 kDa
CC form of the Pseudomonas aeruginosa exoenzyme S (exoS), which can be used
CC to detect, and in vaccines against P. aeruginosa infection. The protein
CC can also be used to produce antibodies, which can be used to detect and
CC treat P. aeruginosa infection, and modify ras protein function in
CC mammalian carcinomas. The exoS gene can be used as a probe to detect P.
CC aeruginosa infection, and in a vector for recombinant exoS production
XX
SQ Sequence 17 BP; 1 A; 6 C; 3 G; 2 T; 0 U; 5 Other;

Query Match 59.0%; Score 11.8; DB 2; Length 17;
Best Local Similarity 66.7%; Pred. No. 6.3e+04;
Matches 10; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 2 GGAGCGCGCGCTCGG 16
Db :|||:|:|:|:|:|:|
15 RGAGCGCGCTCGG 1

RESULT 14
AAH23296
ID AAH23296 standard; DNA; 19 BP.
XX
AC AAH23296;
XX
DT 17-SEP-2001 (first entry)
XX
DE Coding strand of binding site bsG.
XX
KW Nucleic acid binding polypeptide; repressor domain; cardiant; nootropic;
KW circulatory active; anti-inflammatory; dermatological; neuroprotective;
KW cerebroprotective; antibacterial; antifungal; antiviral; antirheumatic;
KW osteopathic; gene therapy; zinc finger; binding site; ss.
XX
OS Synthetic.
XX
XX WO200153480-A1.
XX
XX 26-JUL-2001.
XX
XX 19-JAN-2001; 2001WO-GB0000202.
XX
XX 24-JAN-2000; 2000GB-00001582.
XX
XX 30-MAY-2000; 2000GB-00013102.
XX
XX 30-MAY-2000; 2000GB-00013103.
XX
XX 30-MAY-2000; 2000GB-00013104.
XX
XX (GEND-) GENDAQ LTD.
XX
XX Choo Y, Klug A, Moore M;
XX
XX WPI; 2001-451906/48.
XX
XX

PT Nucleic acid binding polypeptide, used to identify nucleic acids and
PT treat inflammatory, neurological, and dermatological disease, comprises a
PT repressor domain and several nucleic acid binding domains linked by non-
XX canonical linker(s).
XX
XX Example 9; Page 66; 142pp; English.
XX
XX The invention relates to a nucleic acid (NA) binding polypeptide (I)
XX comprising a repressor domain and several NA binding domains (BDs) linked
XX by at least one non-canonical linker. (I) may be used to identify NAs in
XX a complex mixture, to differentiate single base pair changes in NAs, in
XX the manufacture of chimeric restriction enzymes, to produce knock out
XX organisms, and in the treatment of diseases such as: cardiovascular,
XX inflammatory, metabolic, infectious, neurological, rheumatological,
XX genetic, dermatological, and musculoskeletal diseases. The invented
XX methods are used to produce novel NA binding polynucleotides and to
XX modify existing NA binding polynucleotides comprising several NA BDs. The
XX novel polypeptide comprises several nucleic acid binding domains linked
XX by linker sequences. The invented polypeptide is therefore able to span
XX longer or variable gaps, and a greater number of gaps, between DNA
XX binding subsites. Sequences AAH23290-296 represent coding strand
XX sequences of the binding sites used in gel shift experiments with
XX peptides containing flexible linkers
XX
SQ Sequence 19 BP; 1 A; 5 C; 11 G; 2 T; 0 U; 0 Other;

Query Match 59.0%; Score 11.8; DB 4; Length 19;
Best Local Similarity 86.7%; Pred. No. 6.3e+04;
Matches 13; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GGAGCGCGCGCTCGG 16
Db |||||:|:|:|:|:|:|
3 GGAGCGCGCGCTCGG 17

RESULT 15
AAH20644/c
ID AAH20644 standard; DNA; 20 BP.
XX
XX AAH20644;
XX
DT 13-AUG-2001 (first entry)
XX
DE Human telomeric repeat binding factor 2 oligonucleotide 111372.
XX
KW Antisense; phosphorothioate; human; telomeric repeat binding factor 2;
KW inhibitor; premature aging; hyperproliferative disorder; cancer;
KW cyostatic; ss.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX modified_base 1..20 b
XX /mod_base= OTHER
XX /note= "phosphorothioate backbone"
XX modified_base 1..3 a
XX /mod_base= OTHER
XX modified_base 13..20 c
XX /mod_base= OTHER
XX /note= "2-O-methoxyethyl"
XX modified_base 13..20 c
XX /mod_base= OTHER
XX /note= "2-O-methoxyethyl"
XX
XX WO200143752-A1.
XX
XX 21-JUN-2001.
XX
XX 14-DEC-2000; 2000WO-US033954.
XX
XX 17-DEC-1999; 99US-00467642.
XX
XX
```

PA (ISIS-) ISIS PHARM INC.
 XX
 PI Monia BP, Cowsert LM;
 XX
 DR WPI; 2001-398071/42.
 XX
 PT Antisense compounds targeted to nucleic acid encoding telomeric repeat
 PT binding factor 2 useful for treating conditions such as premature aging
 PT and diseases such as cancer.
 XX
 PS Claim 3; Page 80; 108pp; English.
 XX
 CC This invention describes a novel antisense compound (I) 8-30 nucleobases
 CC in length targeted to a polynucleotide encoding human telomeric repeat
 CC binding factor 2 (II) which specifically hybridizes with, and inhibits
 CC the expression of (II). (I) is useful for treating a human having a
 CC disease or condition associated with (II) such as premature aging or a
 CC hyperproliferative disorder especially cancer, by inhibiting the
 CC expression of (II) in human cells or tissues. (I) is useful for
 CC diagnostics, therapeutics, prophylaxis and as research reagents and kits.
 CC The products of the invention have cytostatic activity. This sequence
 CC represents an antisense oligonucleotide used to illustrate the method of
 CC the invention
 XX
 SQ Sequence 20 BP; 1 A; 8 C; 7 G; 4 T; 0 U; 0 Other;
 Query Match 59.0%; Score 11.8; DB 5; Length 20;
 Best Local Similarity 86.7%; Fred. No. 6.3e+04;
 Matches 13; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 5 CGCGGCGCTCGTCA 19
 ||||| |||||
 Db 18 CGCGGCGATCGACA 4
 Search completed: May 3, 2005, 19:48:37
 Job time : 429 secs